View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Texas A&M Repository

CHARACTERIZATION OF CANDIDATE GENES IN ENGLISH COCKER SPANIEL HEREDITARY NEPHRITIS

A Dissertation

by

ZENAIDO TRES CAMACHO

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

December 2003

Major Subject: Veterinary Microbiology

CHARACTERIZATION OF CANDIDATE GENES IN ENGLISH COCKER SPANIEL HEREDITARY NEPHRITIS

A Dissertation

by

ZENAIDO TRES CAMACHO

Submitted to Texas A&M University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Approved as to style and content by:

Joe W. Templeton (Chair of Committee)	David L. Busbee (Member)
George E. Lees (Member)	James E. Womack (Member)
Ann B. Kier (Head of Department)	

December 2003

Major Subject: Veterinary Microbiology

ABSTRACT

Characterization of Candidate Genes in English Cocker Spaniel Hereditary Nephritis. (December 2003)

Zenaido Tres Camacho, B.S., The University of Texas at Austin

Chair of Advisory Committee: Dr. Joe W. Templeton

Six different isoforms of Type IV collagen (colIVα1-6) have been identified. The individual isoforms of colIV are termed alpha chains and are translated from six different COLIV genes (COLIVA1-A6). Collagen Type IV gene products compose the structural framework of basement membranes. The glomerular basement membrane (GBM) is a specialized basement membrane involved in the ultrafiltration processes of the kidney. The colIVα1-α5 chains are expressed in the human GBM while the colIV α1-α6 chains are expressed in the canine GBM. Many inherited diseases of the kidney have been reported and mutations in genes regulating kidney function have been identified. Alport syndrome (AS) is the most common form of human hereditary nephritis (HN). AS is defined as an inherited progressive kidney disorder associated with sensoneural deafness and is characterized by extensive thickening and multilamminar splitting of the GBM when examined by electron microscopy. AS has both X-linked (XLAS) and autosomal (ARAS) modes of inheritance. Mutations in the COLIVA5 gene are responsible for XLAS. A form of HN with characteristic splitting of the GBM with X-linked inheritance has been described in Samoyed dogs. A specific mutation in the

COLIVA5 gene has been identified in Samoyed dogs affected with HN. Mutations in the COLIVA3 and COLIVA4 genes are responsible for ARAS. A form of HN has been identified in English cocker spaniel dogs (ECS) that has been described as autosomal in inheritance and includes GBM abnormalities including extensive lammination characteristic of ARAS. Both ARAS and ECS-HN show loss of the colIVA3 and colIVA4 chains in the GBM when examined with monoclonal anitibodies. ECS-HN has been hypothesized to have the same molecular basis of disease as ARAS. As such, we have isolated and characterized canine COLIVA3 and COLIVA4 sequences from normal dogs and ECS dogs affected with HN and compared the coding regions of these candidate genes.

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

CHAPTER I

INTRODUCTION

The Collagens

Collagens are found in all vertebrate species as well as in Drosophila, sea urchins, nematodes and sponges (Blumberg et al., 1988). It is suggested that all collagens evolved from a common ancestral unit of a 54 base pair (bp) repeat (Chu et al., 1984). Although diverse in their distribution and function, all collagens share features that group them as a family. In 1986 a conference was convened to discuss the biochemistry of collgens (Uitto et al., 1986). At the time of the conference 10 members of the collagen family had been described. Five years previous, only five members of the collagen family had been reported. Whereas many of the initial collagens were described on a biochemical basis of classification, many of the new collagens being reported were based on genetic sequence homologies. One of the main objectives of the conference was to derive a consensus definition of a collagen. As a result, a collagen is described as a triple helical domain that is expressed extracellularly where it performs a structural function. The word collagen is derived from the ancient Greek word kolla for glue and the helical nature of collagens has been known for some time.

This dissertation follows the format of Journal of Cell Biology.

X-ray crystallography was used to illucidate the double helical nature of DNA and this method was also used to determine the triple helical nature of collagen fibers. In 1954, Ramachandran and Kartha first used this technique to identify the interchain hydrogen bonding and the importance of glycine in the amino acid composition of collagen (Ramachandran, 1954). In 1955, they later reported a more detailed spacial analysis, emphasized the importance of proline residues and described the coiled coil nature of collagen (Ramachandran, 1955). In the same year Rich and Crick, using the same methodology, added information on the primary amino acid sequence and proposed a refined structure (Rich, 1955). The helical structure of collagens is essential for the association of the three alpha chains and is termed the collagenous domain of the molecule. The helical structure is made possible by the incorporation of the amino acid glycine at every third residue creating the characteristic collagenous sequence. Collagens are triple helical molecules with the sequence Gly-X-Y where Gly represents the amino acid glycine while X and Y can represent any other amino acid. Most commonly X and Y are represented by the amino acids proline or hydroxyproline (which account for approximately 20% of the collagenous structure) or hydroxylysine which is used to create interchain stabilizing bonds with other hydroxylysine amino acids to stabilize the quaternary structure of the molecule.

The collagens undergo extensive post-translational modification. Lysine is modified by lysyl hydroxylase and the altered amino acid is then used to form interchain crosslinks used for quaternary stabilization. This enzyme was first isolated in 1958 and identified as a necessary physiological component of crosslinking collagens and elastin,

a molecule found in the basement membranes of the heart (Pinnell and Martin, 1968). The necessary nature of these crosslinks is evident by the addition of lathyrogen which prevents crosslinkage. Lathyrogens, such as β-aminoproionitrile (BAPN), are found in plants and when digested by livestock can lead to extreme illness or death. The inherited connective tissue disease Ehlos-Dandrome Syndrome type VI is also associated with a defect in lysyl hydroxilase (Pinnell et al., 1972). Hydroxylysine is also used as a site for glycosylation (Spiro, 1967a). Type IV collagen of the GBM was also shown to be glycosylated (Spiro, 1967b). The importance of the glycosylation is unknown but could possibly help the molecules become more hydrophillic, an essential characteristic of basement membranes. The collagens are synthesized intracellularly in an immature form and then transported extracellulary. The immature collagen is termed procollagen and was first identified in cultured human fibroblasts using chromatographic techniques (Layman et al., 1971) and the bones of rats using pulse-chase experiments (Bellamy and Bornstein, 1971). These reports showed that the conversion from a pre-procollagen to a procollagen occurred extracellularly by an excreted enzyme (then termed a zymogen) that had proteolytic activity. In 1979 the specific proteolytic enzyme responsible for the release of the N- and C-terminal pro-peptides was identified (Kao et al., 1979).

Glycine is necessary at every third residue in the collagenous sequence because its single hydrogen atom as its side chain is the smallest of all amino acids and causes the least possible amount of steric hindrance. Substitution of any other amino acid creates a steric hindrance that disrupts the helical structure. Thus, a base change creating a single amino acid substitution could create disturbances in the structural integrity of

the collagen molecule. This was first seen in mutations in collagen type I genes in osteogenisis imperfecta (Vuorio and de Crombrugghe, 1990). Collagen molecules consist of three monomers wound together in the left-handed triple helix. The individual monomeric molecules are termed alpha chains. The collagens are an extended family of proteins and are categorized as types based on various biochemical techniques (Uitto et al., 1986). Each type of collagen is given a Roman numeral as its type designation (type I collagen, type II collagen, etc.) and at least 19 types of vertebrate collagen have been identified (Kivirikko and Myllyla, 1987; Prockop and Kivirikko, 1995; van der Rest and Garrone, 1991; Vuorio and de Crombrugghe, 1990).

Although the first collagen was isolated from the tail tendon of rats by acid extraction in the widely reported experiments by Nagoette beginning in 1927, the next identified collagen was not identified until 1969 from chicken cartilage (Miller and Matukas, 1969). This second identified collagen created the need for a system of nomenclature. The present nomenclature of collagens categorizes this collagen as collagen type II and the first collagen isolated by Nagoette as collagen type I. In 1971, Kefalides isolated a unique collagen from the anterior lens capsule from bovines and identified this as type IV collagen (Kefalides, 1971). The technique of cyanogen bromide cleavage of purified collagen extract used by Kefalides was the same method used to identify other collagens including the identification of collagen type III isolated from chicken skin a year earlier (Miller et al., 1971). In 1963, with the use of ion exchange chromatography, Piez first identified the alpha chain subunits of collagen type I (Piez, 1963). He termed the most abundant of the alpha chains $\alpha 1$ and the second most

abundant alpha chain α2 and described the nature of collagen type I as being composed of two α1 chains and one α2 chain, thus describing a trimer of $(\alpha1)_2(\alpha2)_1$ (Piez, 1963). A collagen molecule could be composed of a single species of identical alpha chains or a combination of different alpha chain species. A single alpha chain species is associated with a single collagen type. The identity of the alpha chain is defined by the type of collagen with which it is associated and by its order of identity with respect to the other monomers of the same type. For example, the alpha 1 chain of type I collagen (colIVα1) was defined before the alpha 2 chain of type I collagen (colIVα2) and the alpha 1 chains of type I collagen (colIα1) are not the same as the alpha 1 chains of type II collagen (ColIIα1). The role of the collagens is structural and together they are the most abundant protein in the body.

Type IV Collagen

Six different isoforms of mammalian type IV collagen (colIVα1-6) have been identified. The individual isoforms of colIV are termed alpha chains and are translated from six different colIV genes (COLIVA1-A6) (Hudson et al., 1993). All colIV isoforms have a similar function, molecular structure and genetic composition, but different molecular organization and patterns of expression. The characteristics of collagen type IV will be the focus of this section.

Structurally, collagen type IV is an essential component of basement membrane architecture. Type IV collagen allows the formation of a network of collagen molecules rather than a collagen fiber, because the mature colIV molecule has characteristics that make it different from the fiber forming collagens. Type IV collagen contains more than 20 interruptions in the Gly-X-Y collagenous sequence resulting in a more flexible collagen molecule (Brazel et al., 1988; Dolz et al., 1988; Hofmann et al., 1984). These inturruptions in the collagenous sequences range from 21 interruptions for the α1 chain of type IV collagen (colIVα1) (Brazel et al., 1987) to 26 interruptions for the colIVα4 chain (Mariyama et al., 1994). The colIVα2 and colIVα3 chains each have 23 interruptions in their collagenous sequence (Brazel et al., 1988; Leinonen et al., 1994) while the colIVα5 and colIVα6 chains have 22 and 25 interruptions (Pihlajaniemi et al., 1990; Zhou et al., 1994b). The multiple interruptions of collagen type IV are different with respect to the fiber forming collagens like collagen type I which contains an uninterrupted collagen domain (Hofmann et al., 1984).

Terminal non-collagenous sequences of fiber forming collagens are cleaved, but for type IV collagen these sequences are preserved and used for self assembled networks. Three alpha chains of colIV molecules self associate in parallel fashion with their c-terminals at one end of the triple helix and their amino terminals at the other end of the molecule (Timpl et al., 1981). Both the c-terminal globular domain and the amino terminal collagenous domain of type IV collagen are involved in higher order assembly. Each alpha chain has a c-terminal non-collagenous globular domain and collectively the c-terminals of a colIV molecule are termed the NC-1 domain (Timpl et al., 1981). NC-1

domains from two triple-helical colIV molecules associate to form c-terminal linked dimers. Amino terminal sequences from four colIV molecules associate and, collectively, forms a structure termed the 7S domain (Kuhn et al., 1981).

The 7-S domain of type IV collagen was identified from multiple sources including EHS tumor cells, bovine lens capsule and human kidney (Dixit et al., 1981; Risteli et al., 1981). This novel collagen domain had a sedimentation coefficient of 7.2S and resisted bacterial collagenase but had an amino acid sequence similar to basement membrane collagen. The molecule was found to be highly glycosylated, but the origin of the molecule was unknown. The 7S domain is secured with disulfide bonds and hydrophobic interactions (Siebold et al., 1987). The 7S domain has many important features. These include a 21 residue N-terminal non-triple helical region (NH-1), a 117 residue triple helical region (TH-1) that is responsible for the aggregation of the four molecules and a 10 residue non-triple helical region (NH2) that creates an area responsible for flexibility in this region (Siebold et al., 1987). Through these associations, colIV molecules create a network and form the structural foundation of basement membranes.

The network model for the organization of type IV collagen molecules in basement membranes was first proposed after the identification of the non-collagenous c-terminal domain termed the NC-1 domain (Timpl et al., 1981). This model suggested a regular network of type IV collagen molecules generated by sites interacting at the opposite ends of each molecule. By rotary shadowing electron microscopy, it was observed that the 7S domain connected four molecules while the NC-1 domain

connected two molecules and formed a lattice network. The proposed structure indicated a network similar to chicken wire that was continuous and could explain the mechanical strength and flexibility of basement membranes. Later, it was shown by electron microscopy that type IV collagen could self assemble *in vitro* when purified collagen type IV isolated from EHS cells was heated and allowed to reconstitute (Yurchenco and Furthmayr, 1984). This report suggested a more complex network than the initial chicken wire model and indicated lateral assembly of the molecules as opposed to a strict head to head and tail to tail configuration. Lateral associations were shown to exist through similar methods of experimentation and identified the mechanism of lateral associations to be from interactions of the NC-1 domain with specific intermolecular sites (Tsilibary and Charonis, 1986). The lateral association was inhibited by anti NC-1 antibodies. This is an important finding with regards to this research because it showed that, while the NC-1 domain is important for molecular stability, there are several sites within the molecule for investigation with regards to disruption of the matrix that might result in pathogenesis. In other words, the NC-1 domain must have sites internally in associated molecules with which to bind. Further complexity of the matrix assembly was shown *in situ* by the observation of a super helical formation of different monomeric trimers of collagen type IV molecules around an established basement membrane matrix. This modified model of the structural framework of basement membrane matrices indicated a structure much more like a fabric than chicken wire (Yurchenco and Ruben, 1987).

The most common network of colIV molecules in basement membranes is composed of two colIV alpha 1 chains and one colIV alpha 2 chain (colIV(a1)2(a2)1) (Hudson et al., 1993). It was shown that these trimers are formed beginning at the NC-1 domain of the C-terminus, proceeds to the N-terminus 7S region, that the process can be reversed at 50°C, and the trimer can be reconstituted at 25°C (Dolz et al., 1988). Removal of the NC-1 domain with pepsin leads to a loss of the refolding ability. The colIV (α1)2(α2)1 trimer has been found to be present in most basement membranes and has wide tissue distribution (Trueb et al., 1982). The minor species of collagen type IV (α3,α4, α5, and α6), have been shown to have a limited tissue distribution (Kleppel et al., 1989a). These six colIV alpha chains could in theory form 56 different kinds of triple helical monomers (Hudson et al., 1993). Definitive evidence regarding all of the molecular associations of the colIV alpha chains is still an active area of research, although some molecular associations have been reported. As previously mentioned, the colIV (α1)2(α2)1 trimer has been identified in most basement membranes. Using co-precipitation techniques with monoclonal antibodies to colIV NC-1 domains, a heterodimeric complex of two colIVα3 and one colIVα4 chain in the GBM has been reported (Johansson et al., 1992). Earlier studies in the same tissue with the same techniques indicated a homotrimeric structure of three colIV α3 chains (Saus et al., 1988). ColIVα3 homodimers, (colIVα3)3, have also been reported in embryonic tissues (Haralson et al., 1985). More recently, investigators have reported a disulfide crosslinked complex of colIVα3-α4-α5 chains of type IV collagen from bovine kidney (Gunwar et al., 1998). This finding could explain why patients with Alport Syndrome frequently

show no immunohistochemical staining of either the colIVα3, α4 or α5 chains which are normally present in human GBM (Kashtan et al., 1996; Kleppel et al., 1989a; Nakanishi et al., 1994; Yoshioka et al., 1994). In these patients, genetic analysis shows that the molecular basis of disease has been identified as a mutation in the COLIVA5 gene and that the COLIVA3 and COLIVA4 genes are unaffected and transcribed at normal levels. Therefore, these results suggest that mutations in the gene encoding the colIVα5 chain causes defective assembly of not only the colIVα5 chains, but also the colIVα3 and colIVα4 chains in the GBM. For our studies the converse could also be true, mainly, that mutations in either the COLIVA3 or COLIVA4 genes that disrupt protein structure could also disrupt the assembly of colIVα3-α4-α5 trimers and thus lead to the same GBM defect. The exact nature of the colIV interactions is still unknown. The molecular assembly of colIV products could involve multiple combinations the colIVα3, colIVα4, colIVα5 and colIVα6 chains. The association of these chains could also be at the transcription/translation level because, as will be discussed below, mRNA expression of the colIVα3, α4 and α5 chains appears to be coordinated (Thorner et al., 1996). Research on seminiferous tubule basement membrane in bovine testes has shown three major hexameric complexes of colIV molecules (Kahsai et al., 1997). One is the classic colIV$(\alpha 1)_1(\alpha 2)_2$ network, and the others represent two novel networks. One of the novel networks involves all of the colIV alpha chain species (colIVα1-α6), while the other novel network is restricted to the colIVα3-α6 chains. Beyond these novel networks, using monoclonal antibodies and HPLC technology the content of colIVα3 in these

basement membrane structures was calculated to be 40% while the content of the colIVα4 chains was 18% (Kahsai et al., 1997). The result of this research shows highly complex, interactive, coordinated and heterogeneous nature in the composition of colIV molecules in basement membrane matrices.

Collagen type IV genes are highly conserved. Our research using Genbank of comparative sequence analysis of corresponding colIV alpha chains in different mammalian species, consistently show over 80% homology at the nucleotide level. COLIV genes share similar genetic structure and all consist of many exons over large areas of genomic DNA. Each mammalian COLIV gene has over 50 exons and covers at least 100 kb of genomic DNA (Soininen et al., 1989; Zhou et al., 1994a; Zhou et al., 1994b). COLIV genes generate similar sized transcripts of over 5,000 bases and produce similar sized products. The colIVα1 chain represents the smallest product with 1,642 amino acids (Soininen et al., 1987) and the colIVα2 chain is the largest with 1,676 amino acid residues (Hostikka and Tryggvason, 1988). The human col IV α3, α4 α5 and α6 chains have 1,642, 1,652, 1,659 and 1,670 amino acids respectively (Leinonen et al., 1994; Mariyama et al., 1994; Zhou et al., 1992), (Zhou et al., 1994b)).

Although the colIV "classical" chains (colIVα1 and colIVα2) have been shown to be present in most basement membranes, the minor species of colIV (colIVα3, colIVα4, colIVα5 and colIVα6) have a limited tissue distribution. The limited distribution of these products suggests a specific functions in the tissues where they are expressed. The isolation of the minor species of collagen all arose from investigations into the molecular composition of the GBM. As such, much of the investigation into

collagen type IV tissue distribution has been focused primarily on basement membranes of the kidney. Immunohistochemical staining with monoclonal antibodies can be used to study the expression of specific proteins. Chain-specific antibodies directed against individual collagen type IV α chains can be used to detect expression of these products in specific basement membranes.

In the kidney, the colIVα3 and colIVα4 chains are expressed predominately in the GBM and are also detected to a minor extent in Bowman's capsule and the distal tubules (Butkowski et al., 1989; Kleppel et al., 1989b) and, extrarenally, these chains are also detected in the testis (Kahsai et al., 1997). The colIVα5 chain is also detected in the above locations, as well as the collecting ducts of the kidney and, extrarenally, in the skin (Kleppel et al., 1992). The restricted distribution of these molecules suggests a specified role in these structures.

The homology of the collagen gene family with respect to molecular techniques can be seen in the history of the isolation of the COLIVA5 gene (Pihlajaniemi et al., 1990). In an investigation to isolate collagen type V genes (not COLIV genes), a human placental cDNA library was screened with an avian fribrillar collagen gene. This probe was expected to hybridize strongly to collagen type V genes and several colonies with positive signals were isolated. Two of the clones were COLIA1 genes and one was a COLIVA1 sequence. Another clone isolated in this study had interruptions of the collagenous sequence and was identified as a previously unidentified type IV collagen sequence, COLIVA5.

The COLIVA6 gene is X-linked and lies head to head, adjacent to the COLlIVA5 gene, where they share a common promoter (Oohashi et al., 1994; Sugimoto et al., 1994; Zhou et al., 1994a). The isolation of the COLIVA6 gene followed the identification of a distinct pattern found in the COLIV genes. In the analysis of the COLIVA1-A5 sequences, it was found that the genes could be divided into two distinct groupings. Based on DNA sequence homology, the COLIVA1, COLIVA3 and COLIVA5 genes were found to belong to an α1-like class and the COLIVA2 and COLIVA4 genes were classified as α2-like genes. This nomenclature also identified the colIVα1 and colIVα2 as "classical" chains to distinguish them from the other tissue specific chains. Because it appeared that there was a missing member of the alpha 2-like genes, and it was shown that the COLIVA1 and COLA2 genes are arranged in a head to head configuration and shared common promoter elements, it was hypothesized that the COLIVA5 gene was likewise paired with an unidentified collagen type IV gene. Using the homology inherent in these genes, a COLIVA4 clone was used to screen a human lens library, and several overlapping clones were isolated (Oohashi et al., 1994). Characterization of these sequences showed them to contain a novel COLIV isoform that was identified and classified as the COLIVA6 gene (Oohashi et al., 1994). Mapping of COLIVA6 localized the gene to the same region as COLIVA5 on chromosome Xq22 and showed that the genes are also transcribed by a common promoter (Sugimoto et al., 1994). Although COLIVA6 is expressed in some tissues in the kidney such as Bowman's capsule and the collecting ducts, it is not expressed in the GBM of humans or in young dogs, but is expressed in older dogs (Peissel et al., 1995; Lees et al., 2000).

As mentioned previously, the colIVα1 and colIVα2 chains were first identified in 1971 by Kefalides from cyanogen bromide cleavage of bovine anterior lens capsule (Kefalides, 1971). The genes for the colIVα1 and colIVα2 monomers of type IV collagen have been isolated, sequenced and located autosomally on human chromosome 13 (Griffin et al., 1987; Saus et al., 1989; Soininen et al., 1988). These genes in the human are arranged in a head to head configuration and are transcribed by a common bi-directional promoter (Poschl et al., 1988; Soininen et al., 1988). This promoter region had no detectable transcriptional activity in transient gene expression assays after fusion to the chloramphenicol acetylase (CAT) gene in either direction indicating the necessity of additional elements for efficient and tissue specific expression of both genes (Poschl et al., 1988). No molecular defects in these genes have been identified in HN. It is unlikely that mutations in COLIVA1 and COLIVA2 genes would be found in specific diseases since these genes are expressed in embryonic development and their products are found in most basement membranes (Trueb et al., 1982). In other words, structural defects in these molecules would not be tolerated in development and would lead to wide ranging pathologies. In fact, temperature sensitive mutations in the COLIVA1 and COLIVA2 genes in *C. elegans* cause embryonic lethality (Guo 1989). It is suggested that genetic defects resulting in a loss of function in a tissue specific collagen species, such as COLIVA3 or COLIVA4, is more likely in autosomal forms of HN.

The colIVα3 chain was first isolated using antibodies from patients with Goodpasture syndrome (GS). GS is an autoimmune disorder in humans where patients have circulating antibodies to GBM components and develop rapidly progressive

glomerulonephritis. After collagenase digestion, preparations from bovine GBM that reacted with Goodpasture sera identified the NC-1 domain of collagen type IV as the targets of Goodpasture antibodies (Wieslander et al., 1984). Further investigation with lens basement membranes using reverse-phase HPLC, SDS-PAGE and amino acid sequencing showed the Goodpasture antigen to be a novel collagen type IV isoform designated colIVα3 (Butkowski et al., 1987; Saus et al., 1988). Another result of this same investigation was the identification of another novel collagen type IV isoform designated colIVα4. Although the above investigation only briefly characterized these novel proteins with respect to differential electrophoretic migration patterns and immunoreactivity, the discovery of new colIV isoforms opened the door to many other findings including the understanding of basement membrane structure and function. These findings also helped to explain the molecular basis of many inherited basement membrane pathologies that will be discussed below.

The human COLIVA3 and COLIVA4 genes have been mapped to human chromosome 2 and, similar to the genes described above, are situated in a head to head configuration (Kamagata et al., 1992; Mariyama et al., 1992; Morrison et al., 1991). They have also been mapped autosomally in the dog (personal communication). In the human genome, both the COLIVA3 and COLIVA4 genes have been suggested to cover over 100 kb and to contain over 50 exons (Boye et al., 1998; Heidet et al., 2001). Each gene produces a product of approximately 1,700 amino acids from an ~5,000 base coding region (Leinonen et al., 1994; Mariyama et al., 1994). Partial bovine COLIV cDNA sequences have also been isolated using human sequences as primers for PCR

amplification (Morrison et al., 1991). The COLIVA3 and COLIVA4 genes produce major components of the GBM and aberrations in these genes have been found in humans with autosomal HN which will be discussed in later sections (Boye et al., 1998; Lemmink et al., 1994; Mochizuki et al., 1994).

Extracellular Matrices

Extracellular matrices (ECMs) are classified into two categories. One form of ECM fills spaces between cells and has many receptors for the binding of cytokines and growth factors. The other form of ECM creates basement membranes, which are thin sheets around tissues and were initially termed reticulin (Rambourg and Leblond, 1967). Together the separate components of basement membranes create a scaffold for cells and maintain the structural integrity of many tissues. Although having similar components, basement membranes are not necessarily the same in all tissues. Early immunological assays of the canine kidney showed differences between the basement membranes of the tubules and of the GBM (Krakower, Goodman 1955). Their research showed that although these basement membranes were made predominantly from collagen, the antigenic epitopes differed between these structures although, at the time of this report, the collagens causing the immunological reactions they observed had not been identified. The ECM plays an important role in modulating the behavior of cells with which it interacts. The major molecules that constitute the ECM include collagen type

IV, entactin/nidogen (E/N), laminin and heparine sulfate glycoprotein (HSGP). A model system for the study of ECM components is the Engelbreth-Holm-Swarm (EHS) cell line, which is derived from a benign mouse tumor and produces abundant ECM molecules (Swarm, 1964). The EHS matrix consists of 60% laminin, 30% colIV, 3% heparine sulfate proteoglycans and 6% E/N (Kleinman et al., 1986).

The glomerular basement membrane (GBM) is a specialized ECM created predominately by podocytes of the kidney. Aberrations in the GBM have been shown to be associated with podocyte abnormalities including effacement and resulting kidney pathologies including proteinuria. The major components of the GBM and their molecular interactions will be discussed in this section, except for colIV whose role in the GBM has been described above and will also be discussed in subsequent sections.

Although constituting a small proportion of overall basement membrane mass, E/N will be discussed first because of its ability to bind to all of the other major components of the GBM. Two groups characterized this molecule independently from different sources. The first group described a basement membrane component isolated from a mouse carcinoma cell line and named the molecule entactin (Carlin et al., 1981). Another group isolated a similar product from the EHS matrix and named the molecule nidogen (Timpl et al., 1983). Both molecules were glycoproteins of approximately 150 kD. Amino acid and cDNA sequence analysis of entactin and nidogen in mouse and human demonstrated that both were identical molecules and thus can be referred to collectively as E/N (Durkin et al., 1988; Mann et al., 1989; Nagayoshi et al., 1989). The gene for E/N has been located autosomally on human chromosome 1 (Olsen et al.,

1989), and it has been shown that E/N expression can be detected at the 8-16 cell stage in mouse development (Dziadek and Timpl, 1985; Wu et al., 1983). Expression of E/N has also been found in the GBM (Saxena et al., 1990). Electron microscopic rotary shadowing of E/N shows a linear molecule composed of three globular domains (G1, G2 and G3) separated by two rod-like domains, and has a total length of approximately 30nm (Fox et al., 1991). E/N binds to laminin via the G3 domain (Mann et al., 1988) and to collagen IV via its G2 domain (Aumailley et al., 1989; Fox et al., 1991). E/N has also been shown to bind to the protein core of HSGPs in basement membranes (Battaglia et al., 1992). Thus, although E/N was first identified by co-immunoprecipitation with laminin, it appears that apart from being able to bind to itself, E/N can bind to all of the other major basement membrane components. Its ubiquitous and early expression indicate a strong role in basement membrane assembly.

Other isoforms of E/N have been reported in mouse and human (Kimura et al., 1998; Kohfeldt et al., 1998). The mouse isoform has been termed entactin-2 while the human isoform is termed nidogen-2. These new isoforms can be conjointly termed E/N-2. These molecules show similar structure to the original E/N-1 molecules, but share less than 50% sequence homology to the E/N-1 isoforms (Kohfeldt et al., 1998). E/N-2 is proposed to have similar function to E/N-1 because E/N-2 binds to laminin, but at a different site and with less affinity than E/N-1, and has colIV and HSGPs binding ability equal to the E/N-1 isoforms (Erickson and Couchman, 2000). The discovery of E/N-2 suggests the other E/N isoforms may exist and may be a growing family of ECM molecules with different tissue specific functions.

Many non-collagenous glycoproteins are building blocks of ECMs, but the laminins are the most ubiquitous and abundant in basement membranes. Laminin was first isolated from an EHS matrix (Timpl et al., 1979). Laminins are large heterotrimereric molecules of 400-600 kD and it has been shown that the laminins can self assemble *in vitro* into independent networks via intermolecular bonding (Yurchenco et al., 1992; Yurchenco et al., 1985). The network forming ability of laminin is thought to be important for basement membrane stability. The laminins have evolved into a family of proteins with many different isoforms. These laminin isoforms include five different α-chains (α1-α5), three different β-chains (β1-β3) and three different γ-chains (γ1-γ3) (Burgeson et al., 1994). The composition of a mature laminin molecule includes one of each of the three chains. It is possible to predict which laminin trimers are present in basement membranes by co-localization with immunohistochemical staining of individual laminin chains. Theoretically, 45 trimeric formations of laminin chains can exist (5α chains x 3β chains x 3γ chains). To date, however, only 12 different laminin isoforms have been reported (Burgeson et al., 1994). The most abundant laminin in the GBM is laminin-11 composed of a trimer of α5β2γ1 chains (Miner et al., 1997). Laminin-9 (α4β2γ1) has been reported in developing GBM (Miner et al., 1997), but the α4 chain gradually disappears as the glomerulus matures (Miner et al., 1998). The functions of the laminins have been inferred by identification of mutations in these genes in specific diseases, or by knockout experiments of specific laminin chains in mice. It is presumed that the knockout of a specific laminin chain will result in the loss of function of the mature laminin trimeric molecule. Knockout experiments involving the β2 chain

in mice results in early death due to massive proteinurea and show fusion of the podocyte foot processes, although, the GBM appears normal (Noakes et al., 1995). While the α5 chain of laminin is also present in laminin-10 (α5β1γ1) which is present in most basement membranes of the kidney (Miner et al., 1997), laminin α5 genetic knockouts also affects laminin-11 (α5β2γ1) in the GBM. As expected, the α5 knockouts have much more devastating effects on kidney development and the mice die *in utero* (Miner et al., 1998). The α5 chain of laminin has been shown to be essential for glomerulogenisis, and the laminin α5 knockout mice also have neural tube defects that indicate the importance of this chain in other tissues (Miner and Li, 2000).

Like the E/N isoforms described above, the laminins can form associations with other ECM molecules. A strong component of laminin function in basement membrane architecture involves molecular associations with integrins. Integrins are a diverse family of integral membrane glycoproteins that function to mediate interactions of cells with the ECM. Integrins also function to mediate cell:cell interactions and cell signaling (DeSimone et al., 1987). Integrins are dimeric proteins that contain α and β chains. There have been 19 integrin α and 8 integrin β chains reported and 25 different combinations of these chains have been described (Humphries, 2000). The molecular association of laminins and integrins directly links the ECM to the surface of the cells. Many regions of laminin have been reported to associate with integrins and many integrin–laminin associations have been reported (Belkin and Stepp, 2000).

As mentioned previously, the laminins bind E/N in a 1:1 equimolar association; that is, one molecule of laminin binds one molecule of E/N. Unlike E/N, however,

laminin can self assemble to form an independent network in basement membranes (Yurchenco et al., 1992). The specific epitope of laminin that binds E/N has been located on the γ chains with the γ1 chain having the strongest association, and antibodies to this region inhibit laminin:E/N binding (Mayer et al., 1993). Thus, the laminin-E/N interaction connects the colIV and laminin networks as well as other ECM components such as the HSPGs. The direct binding of laminin to HSGPs can increase these interactions. The triple helical core of laminin has been shown to bind to agrin, a specific HSGP in the in the GBM (Denzer et al., 1997).

The carbohydrate content of basement membranes was first demonstrated by periodic acid-shiff staining techniques and attributed to glycoproteins in 1957 (Leblond and Inoue, 1989). Heparine sulfate glycoproteins (HSPGs) are found in all basement membranes and consist of a protein core to which several carbohydrate chains are attached (Paulsson et al., 1987). HSPGs contain heparine sulfate as its glycosaminoglycan (GAG) side chain that is composed of unbranching repeating disaccharide units with the aminosugar glucosamine. The glucosamine component of HSPG is extremely sulfated and creates a hydrophilic positively charged environment in basement membranes (Kanwar and Farquhar, 1979b). The anionic nature of the GBM created by the presence of HSPG is extremely important to GBM permselectivity and will be discussed in later sections.

Based on immunoflourecent studies with antibodies directed against the core proteins of HSGPs, the most abundant HSGP in the GBM is agrin (Groffen et al., 1997; Groffen et al., 1998). Low levels of agrin are also detected in other basement membranes

in the kidney in the rat, but are attributed to different agrin isoforms (Raats et al., 1998). The 95nm protein core of agrin has many domains but its most N-terminal region has been shown to be necessary for the binding of laminin-1 (Denzer et al., 1997; Denzer et al., 1995). Agrin has also been shown to bind laminin-2 (Cotman et al., 1999) and integrin $\alpha 5\beta 1$ (Martin and Sanes, 1997) and, although these specific isoforms have not been identified in the GBM, as has been discussed, intermolecular associations are essential for GBM architecture. Although this investigation focuses on the components of the GBM, agrin is also known as a key component of the neuromuscular junction (Denzer et al., 1998). Agrin was first identified as a synaptic organizing protein associated with aggregation of acetylcholine receptors and acetylcholinesterase at regenerating neuromuscular junctions *in vivo* in Torpedo (an oceanic ray) (Nitkin et al., 1987). Initially identified as an aggregating factor, the molecule was termed agrin. Agrin has been characterized as a large glycoprotein of over 400 kD when fully glycosylated (Tsen et al., 1995) and has been mapped autosomally in both human and mouse (Kallunki et al., 1991; Rupp et al., 1991). The importance of agrin has been demonstrated with knockout experiments in mice which result in aberrant kidney development and death at birth (Gautam et al., 1996).

Other proteoglycans in basement membranes of the kidney include perlecan, bamacan and colXVIII. Perlecan is widely expressed in the kidney, but has restricted distribution and is not predominant in the GBM (Groffen et al., 1998). Bamacan is a chondroitin sulfate glycoprotein which also has widespread distribution in the kidney, but has not been detected in the GBM (Wu and Couchman, 1997). ColXVIII has been

shown to be a glycoprotein (Halfter et al., 1998) and expression of colXVIIIα1 has been reported in the GBM (Muragaki et al., 1995). ColXVIII has been classified into a subfamily of non-fibrillar collagens with multiple interrupted triple helical domains, but any role in any pathology has yet to be determined (Oh et al., 1994).

One of the fascinating aspects of basement membrane biology is the ability of the individual components to self assemble *in vitro*. Individual basement membrane components from EHS cells (including laminin, E/N, collagen type IV and HSPG) were isolated from purified extracts and, when combined, at physiologic temperatures, spontaneously assembled into structures that resembled the lamina densa of basement membranes (Grant et al., 1989).

The increased knowledge of basement membranes, including the above studies, indicates a complexity of basement membrane architecture through specific molecular associations that reflect the diverse nature and function of all ECMs. For example, it is now problematic to discern the exact point of transition from the ECM, the cell membrane and the cytoskeleton. The specific focus of this study is directed toward the GBM, a specialized ECM. The GBM is subjected to tremendous forces and constant pressure from birth until death. The delicate nature of this structure can be revealed if any component has a loss of function from either environmental influences or genetic aberrations. The composition of basement membrane structure that has been described above is based on specific intramolecular and intermolecular associations and has been focused primarily on molecules expressed in the GBM. Interruptions in these molecular

associations result in decreased performance of the GBM and, as will be discussed below, an overall loss of kidney function.

Kidney Structure and Function

The kidneys regulate arterial pressure, extracellular fluid osmolarity, blood and extracellular fluid volumes, acid-base balances in the body and the excretion of nitrogenous waste products. Although the kidneys perform many physiological functions, the focus of this investigation will concentrate on the kidney as the site for the ultrafiltration of plasma.

Many aspects of human kidney function can be found in numerous textbooks on physiology including “The Textbook of Medical Physiology” (Guyton, 1991) and “Physiology” [Sperelakis, 1996]. Other resources written with specific reference to kidney function include “(The) Kidney” [Brenner 1991] and “Pathology of the Kidney” [Heppinstall, 1992]. These references will be used to describe the basic nature of kidney physiology.

Normally, there are both left and right kidneys and each kidney is profused with one renal artery and drained with one renal vein. In between, the flow from the renal artery is divided into individual arterioles, many of which introduce blood to individual nephrons. Although there are two kinds of nephrons, the cortical nephrons and the juxtaglomerular nephrons, each nephron is capable of formation of urine. Therefore, although nephrons in mammalian kidneys number in the hundreds of thousands, urine

formation can be described using a single nephron. Each glomerulus, it's associated tubule and vascular structure defines a nephron. Blood flows to the glomerulus, the main filter of the kidney, through the afferent arteriole and exits through the efferent arterial. Every component of blood is passed through the capillary structure of the glomerulus. The afferent arteriole is divided into approximately fifty capillaries in each glomerulus. The capillaries of the glomerulus are uniquely fenestrated, that is, the capillaries of the glomerulus have thousands of small holes averaging from 500 A° to 100 A° in diameter. This allows the blood of the capillary structure of the glomerulus to come into direct contact with the underlying basement membrane (GBM). The GBM is sandwiched between the endothelial capillaries of the glomerulus and the epithelial cells that surround the glomerular capillaries. The epithelial cells are called podocytes. Podocytes are specialized epithelial cells that interdigitate with one another to form the foot processes. The interdigitation of podocytes forms slit diaphragms (also known as filtration slits) in the glomerular epithelium. The podocytes of the glomerulus are surrounded by another epithelial structure known as Bowman's capsule. Bowman's capsule forms the first structure of the urinary space and is the collecting site for the primary filtrate that is created by the ultrafiltration of plasma.

Thus, ulrafiltration of plasma is directed in the glomerulus from the fenestrated capillary across the GBM and through the slit diaphragms created by the epithelial podocytes and into Bowman's capsule. From Bowman's capsule the primary filtrate flows into the tubular structure of the nephron where further processing of the primary filtrate occurs. The active participation of the tubular epithelial cells contribute to the

processing of the primary filtrate and is an essential process for the ultimate formation of urine.

One of the basic functions of urine formation can be described as the excretion of unwanted substances. Unwanted substances in the blood include excesses of sodium, potassium and chloride ions as well as end products of metabolism including urea and creatinine. In as it is a basic function of the nephron to clear the plasma of unwanted substances, the nephron must be selective and not allow cells or large proteins to pass into the urinary space. While the average human has about 40 liters of total body fluid, 180 liters of plasma is filtered per day, creating 14 liters of primary filtrate and ultimately 1.4 liters of urine per day. Therefore, the glomerulus filters a large amount of plasma to create a primary filtrate that flows into the urinary space and then into the tubular portion of the nephron. In the tubular structure, the unwanted substances are passed in the formation of urine, while most of the water and many other solutes are reabsorbed. Most of the soluble components of the blood pass through the filtration mechanisms of the glomerulus and are present in the primary filtrate. Electrolytes are reabsorbed with water according to homeostatic demands. Larger serum proteins are normally not allowed passage into the tubular space. One of the smallest of plasma proteins, as compared to fibrinogen at 340 kD or immunoglobulins at 150-900 kD, is albumin with a molecular weight of 69 kD. Albumin is normally not allowed passage through the GBM into the urinary space. While small amounts of proteins are present in the proximal tubule, with normal kidney function, these proteins are cleared during urine formation. Small proteins that may pass through the ultrafiltration processes undergo

pinocytosis via cells of the proximal tubule. Proteinuria refers to the presence of protein in the urine. Hematuria refers to the presence of erythrocytes or cell structures (cell casts) in the urine. Neither is present in normal urine. A wide range of pathologies, many of which will be described in later sections, are associated with abnormalities in the filtration process.

The structure and function of the mammalian kidney has long been a subject of fascination and experimentation. An experimental model of nephritis was established in 1900 by Lindemann by introducing serum from a guinea pig immunized with rat kidney that became nephrotoxic when injected into rabbits. This experimental model was refined further when sonicated preparations of GBM were used in the initial immunization helping to establish the GBM at the target site for nephrogenic antigens (Rohrbach, 1991). Electron microscopy had been used for many purposes, but it was not until 1948 when a method described by Pease and Baker was used to make fine sectioning of biological tissues that this powerful tool could be used to examine fine tissue structure (Pease and Baker, 1950). Their method allowed an ultrathin sectioning of .05 microns and their survey resulted in many findings. They were able to describe the "brush border" of the proximal tubule, the fenestrated nature of the glomerular endothelium and detailed descriptions of the glomerular basement membrane. They describe the GBM as "pictorially most spectacular". They described the GBM as thin (300-350 nm) and existing as a continuous layer around the fenestrated endothelium that is in direct contact with the epithelial cells of the glomerulus. These studies, among others, have shown the GBM to be a specialized structure in the kidney. The molecular

composition of the GBM has been described above. The role of the GBM with respect to ultrafiltration and the importance of changes in the GBM associated with disease will be discussed below.

Mechanisms of Glomerular Permselectivity

The mechanisms of glomerular permeability have been an object of discussion for some time. In 1960, a seminal study was done by Farquhar et al. describing ferritin transfer across the glomerular capillary wall to explore the mechanisms of glomerular ultrafiltration (Farquhar et al., 1961). Although the structure of the renal glomerulus had been described, little was known about the mechanisms involved in the transport of substances from the capillary to the urinary space. At that time two prevailing hypotheses had been established for glomerular filtration. One was that the glomerular basement membrane acted as the main filter of the kidney since the majority of the electron microscopists had described the GBM as the only continuous layer around the capillary wall. The other hypothesis indicated that the spaces between the foot processes of the epithelial cells acted as slit pores allowing for the free passage of water and small molecules but restricted the passage of proteins. The approach of Farquhar et al to explore the mechanisms of glomerular permeability was to directly observe the pathways of glomerular filtration across the capillary wall by introducing a particulate tracer able to be seen by electron microscopy. Their main conclusion, held for many decades, was

that the glomerular basement membrane was the main filter of the kidney in that there were no identified channels across the foot processes. They observed, however, that the filtration was not perfect and that any small molecules, especially proteins, would be absorbed through pinocytosis in the epithelium. They further concluded that the function of the GBM is probably dependent on both the epithelium and endothelial layers and any agent that would interfere with these cells could affect the chemical composition and therefore the permeability of the basement membrane. This group reinforced their conclusions with a report using graded dextrans to determine the molecular weight limitations of glomerular filtration (Caulfield and Farquhar, 1974). These results found that the size limitations of glomerular filtration using dextran sulfate were 40-50 kD. While 14 kD dextran was cleared from the bloodstream easily, progressive size resulted in decreased clearance until approximately 55-60 kD where there was a sharp decline of clearance. This corresponded to the clearance of albumin, which at a molecular weight of 68 kD does not cross the GBM. In other words, they identified the size limitation of the filtration of the GBM at 10A° while the size of albumin is 36A°. Their results helped to reinforce that the GBM is the primary filtration barrier to plasma proteins. Thus, the size and weight of molecules is important for glomerular permselectivity and further studies showed that the charge of a molecule is also significant in determining “charge-selectivity” (Chang et al., 1975). Again, using the normally neutrally charged dextran sulfate in the previous research, charges were added to the molecules to make it either positive or negatively charged. Their results showed that positively charged molecules pass through the GBM much faster than a neutrally charged molecule and that a

negatively charged molecule is inhibited from crossing. It is interesting to note that albumin is a negatively charged molecule. The mechanisms of charge and filtration is based on the fact that the heparine sulfate glycosaminoglycans (GAGs) that are imbedded in the GBM are anionic (Kanwar and Farquhar, 1979a). In fact, the loss of these anionic sites is associated with some forms of nephrosis (Caulfield and Farquhar, 1978; Rohrbach et al., 1983) and the removal of GAGs using GAG-degrading enzymes allows albumin to pass through the GBM (Rosenzweig and Kanwar, 1982). In more recent reports, an acute, selective proteinurea could be induced in rats by a single intraveinous injection of a monoclonal antibody directed against the most abundant HSPG in the GBM, agrin (van den Born et al., 1992).

Some of the first evidence for an alternative view of glomerular ultrafiltration came from studies using peroxidase molecules as tracers (Graham and Karnovsky, 1966). The peroxidase molecules were twice the size of human albumin and large enough to be retained by the glomerular structure. These studies found that a greater concentration of reaction product was retained in the portion of the basement membrane adjacent to the epithelial foot processes. They further reported that much of the product was retained on the epithelial membranes and between the foot processes adjacent to mesangial areas and more specifically in the epithelial slits. Only after some time did the peroxidase product stain the glomerular basement membrane. There was little disagreement that filtrate passed from the capillary lumen through the endothelial fenstrae, across the basement membrane, and through the epithelial slits into the urinary space, but these studies suggested the restriction of passage occurred at the epithelial

slits. This challenged the hypothesis that the GBM was the principle glomerular filter and suggested that the main filter barrier is distal to the GBM located at the epithelial slits; that the GBM was regulated in function as a “course filter” and that the epithelial slits are the principle filtration barrier.

During this early stage of research, however, the structural evidence for a mechanism at the slit pore was unknown, but the use of new staining techniques identified structures in support of this hypothesis. With the use of tannic acid fixative for electron microscopy, a structure termed the “slit diaphragm” was identified at the epithelial slits between the foot processes (Rodewald and Karnovsky, 1974). The slit diaphragm was identified as an isoporous structure with rectangular pores of approximately 40A° across, 140A° through the slit diaphragm and 70A° in length. The dimensions of the pores found in the slit diaphragm were smaller than the 70A° width of albumin, and supported the hypothesis that the slit diaphragm was the principal filtration barrier to plasma proteins. Furthermore, these studies showed that the slit diaphragm was a continuous, two-dimensional band that extended to all foot processes so it is a necessary structure for molecules to cross this in the filtration process and diminished the uniqueness of the GBM in this regard. Detailed analysis of the electo-imagery reported in this same investigation described another new “zipper-like” spacing of cross bridges within the slit diaphragm although the molecular nature of these novel structures could not be revealed until recently.

Mapping the genetic defect in congenital nephrotic syndrome of the Finnish (NPHS1) type, an autosomal recessive disorder in humans, identified the first of these

molecules in the slit diaphragm. Congenital nephrotic syndrome (CNS) is a heterogeneous group of renal diseases that result in increased postnatal glomerular permeability and proteinurea early in life. Mapping of NSPH1 using positional cloning localized the gene to a 150 kb region on chromosome 19q13.1 and sequencing of this region revealed 11 candidate genes, one of which was mutated in affected families (Kestila et al., 1998). Over 30 different mutations in the NSPH1 gene (Nphs1) have been identified (Lenkkeri et al., 1999). Nphs1 codes for nephrin, a large transmembrane glycoprotein, that is 26 kilobases, has 29 exons, and codes for a 4.3 kb mRNA and is a member of the Ig superfamily (Kestila et al., 1998). Rat and mouse nephrin genes have also been characterized and show high homology and similar organization to the human genes (Putaala et al., 2000). Nephrin has been specifically located to the slit diaphragm in human (Ruotsalainen et al., 1999) and mouse (Holzman et al., 1999). It is postulated that pairs of nephrin molecules compose the molecular zipper between the foot processes of podocytes. A model for NPHS1 has been created by inactivation of the Nphs1 gene in mice that will provide a powerful system for future studies (Putaala et al., 2001). Also, the glomerular specific promoter for the nephrin gene has been inserted into two systems that may facilitate the identification of transacting factors that are required for podocyte-specific expression of GBM molecules. (Moeller et al., 2000; Wong et al., 2000)

Another necessary component for slit diaphragm assembly was previously well characterized, and yet its function in the kidney has only recently been revealed. The CD2-associated protein (CD2AP) is critical for stabilizing contacts between T cells and antigen presenting cells and is termed CMS in humans. Knockout mice in this gene were

developed in an effort to determine the loss of CD2-associated protein (CD2AP) on immune function (Shih et al., 1999). A knockout of the first intracellular SRC homology domain (SH3) in blastocysts resulted in a diminished immune capacity, but these mice also developed a form of congenital nephrotic syndrome. The cause of renal failure was determined to be due to the loss of foot processes in the glomerulus. Further, it was shown that the CD2AP co-immunoprecipitated with nephrin (Shih et al., 1999). CD2AP is expressed in the slit diaphragm and binds nephrin via a novel c-terminal domain (Shih et al., 2001). Because of the known function of CD2AP in T cells, it is suggested that CD2AP anchors nephrin to the internal protein fibers that form the podocyte cytoskeleton to help form and stabilize the slit diaphragm (Yuan et al., 2002).

The study of pathologies with apparent filtration barrier defects can provide useful models for the discovery of the genes and molecules responsible for normal glomerular filtration and the molecular basis of disease. This has been demonstrated in the study of focal segmental glomerulosclerosis (FSGS). FSGS is heterogeneous group of disorders manifested as non-specific renal lesions secondary to other disorders or as a specific inherited condition. FSGS may be defined by the presence of segmental sclerosis in some, but not necessarily all glomeruli, and is characterized by increased urinary protein excretion and decreased kidney function leading to end-stage renal failure (Ichikawa and Fogo, 1996).

The steroid-resistant form of FSGS in humans (SRN) is an autosomal recessive disorder and has been mapped to chromosome1q25-31 (Fuchshuber et al., 1995). Recently, mutations in the Nsph2 gene have been reported in families with SNR (Boute

et al., 2000). The product of Nsph2 has been termed podocin because the expression of this gene is restricted to fetal and adult podocytes. The podocin gene codes for a 42 kD protein from a 1,149bp open reading frame. It is a putative integral membrane protein and shows homology to human stromatin, which appears to be an ion channel associated with the cytoskeleton. Although the actual function of podocin is unknown, it is possible that it may be involved in organizing the cytoskeleton of podocyte foot processes at the slit diaphragm and may also interact with other podocyte proteins necessary for slit diaphragm assembly. It has been reported that podocin co-localizes with nephrin to the actin cytoskeleton protruding from the tips of actin filaments (Saleem et al., 2002). The cytoplasmic side of podocin has molecular homology for binding to an SH-3 binding site similar to the structure found in the CD2AP above.

The gene responsible for the autosomal dominant form of FSGS has been mapped to human chromosome 19q13 and mutations in the gene encoding α-actin-4 (ACTN4), which resides at this locus, have been reported in three families (Kaplan et al., 2000). Although the exact function of ACTN4 with regards to FSGS is unknown, expression of this gene occurs in capillaries, glomerular podocytes and in other tissues such as circulating lymphocytes. ACTN4 is an actin crosslinking protein and it is postulated that regulation of the actin cytoskeleton of glomerular podocytes may be altered in FSGS thus interfering in normal filament assembly in the foot processes. This indicates that the actin cytoskeleton in podocytes may be involved in normal glomerular ultrafiltration and in the pathogenesis of proteinurea. Because ACTN4 is also expressed in capillaries, it is possible that a loss of function in this gene could also change the renal

vasculature cytoskeleton and thus alter glomerular hemodynamics leading to the described phenotype. Regions relating to other forms of inherited FSGS have been mapped in the human, but the specific genes for these syndromes have not yet been identified.

These studies show a complex design of glomerular ultrafiltration that includes the coordinate function of the slit diaphragm and the GBM involving many gene products. The function of each of these genes is associated and/or dependent on intermolecular and/or intramolecular associations of other GBM components. The loss of function of any of these components can, therefore, lead to molecular dissociation and subsequent failure of overall glomerular ulatrafiltration processes and, hence, an overall loss of kidney function. As will be discussed, the products of the COLIV genes have been shown to be specifically involved in these processes.

Alport Syndrome and Other Diseases of the GBM

Although there are many diseases that are aimed at various parts of the nephron, like the slit diaphragm pathologies described above, diseases referred to as HN can be located directly to components of the GBM. Owing to its similarities to the form of hereditary nephritis in English cocker spaniel dogs which is the focus of this investigation, Alport Syndrome (AS) will be discussed more fully than other diseases of the GBM.

In humans, AS is the most common form of HN with an estimated gene frequency of 1:5000 (Atkin, 1988). Dr. Cecil Alport in 1927 first reported the symptoms associated with a new clinically relevant syndrome, which is now referred to as Alport Syndrome (AS) (Alport, 1927). Dr. Alport described an inherited nephrotic condition that was associated with deafness where male members of a family progressed to end stage renal disease (ESRD). The characteristics of AS have since been shown to be very heterogeneous. In 1988 AS was classified into six different types based on clinical criteria (Atkin, 1988). Type I was classified as juvenile Alport with deafness, type II was X-linked with deafness, type III was X-linked adult with deafness, type IV was X-linked adult without deafness, type V was autosomal with deafness and thrombocytopathy and type VI was autosomal with deafness.

The most common and identifiable characteristic of AS are renal abnormalities. In individuals with hereditary nephritis, the lamina of the GBM show structural abnormalities including overall thickening and splitting that can be seen by electron microscopy (Churg and Grishman 1972; Hinglais et al. 1972; Spear and Slusser 1972). AS can be diagnosed when the characteristic splitting of the GBM is observed (Heptinstall, 1992).

The most common extra-renal defect associated with AS is deafness, which is frequent, but not universal, and occurs in approximately 50% of all patients (Wester et al., 1995). While renal disorders are not always accompanied with hearing loss in AS, members with hearing loss in families diagnosed with AS always have accompanied renal disorders (Gubler et al., 1981). Also, there have been no reports of deaf males

without kidney disease transmitting AS to any offspring (Gubler et al., 1981). While the exact mechanism for the loss of hearing in AS patients has not yet been determined, the defect has been isolated to areas of the inner ear (cochlea).

Other common extra renal abnormalities associated with AS include a variety of ocular defects. While ocular defects may be present in less than half of all AS cases (Chance and Stanley, 1977), a condition known as anterior lenticonus is virtually pathognmonic of AS (Nielsen, 1978). Anterior lenticonus is the protrusion of the central portion of the lens into the anterior capsule and is most associated with AS families with deafness and with rapid progression to end stage renal disease. Other ocular defects seen in patients with AS include changes in various basement membranes of the eye including Bruch's membrane, Descemet's membrane and corneal basement membranes (Govan, 1983; Rhys et al., 1997; Teekhasaenee et al., 1991).

Other extra renal abnormalities associated with AS include leiomyomatosis (hyperplasia of smooth muscle) (Cochat et al., 1988), and hematological defects such as thrombocytopenia (Epstein et al., 1972). The presence of hematological defects in AS is interesting because unlike the tissues described above, blood cells do not have classic basement membrane architecture.

Since the diagnosis of AS has been based on certain clinical findings, confusion over diagnostic procedures and AS classification is understandable given the heterogeneity described above. Further complications over the proper diagnosis and classification of AS can be attributed to heterogeneity within a particular type of AS. The severity of symptoms varies from person to person and small kindereds contain too

few individuals to meet many of the clinical and histologic criteria for proper classification. Precise diagnosis and classification of AS is important for several reasons including genetic counseling and the establishment of individual prognoses. Furthermore, there are many diseases that can be confused with AS.

Although difficulties in the diagnosis of AS have been reported, a core of traits can be associated with all types of AS that have been described above. These include a hereditary disease with chronic hematuria that is progressive and characterized ultrastructurally by irregular GBM structure. Furthermore, AS can be differentiated from immune related glomerulonephritic diseases. Most importantly, AS can be generally classified as having either X-linked or autosomal inheritance.

Of the many types of AS that have been classified, the molecular basis of disease was first identified for X-linked AS (XLAS). In the initial investigation of the molecular basis of XLAS, it was known that basement membrane ultrastructural aberrations were pathognomonic of AS and that collagen type IV made up the structural framework of the GBM (Churg and Sherman, 1973; Hinglais et al., 1972; Spear and Slusser, 1972). Also, the genetic element responsible for XLAS had been mapped to the X chromosome (Atkin et al., 1988; Szpiro-Tapia et al., 1988). The identification of the minor species of type IV collagen including the COLIVA5 gene (Hostikka et al., 1990; Pihlajaniemi et al., 1990) and the mapping of COLIVA5 to the X-chromosome (Hostikka et al., 1990) identified COLIVA5 as a strong candidate gene. These reports focussed efforts toward the screening of this gene in families with X-linked AS. The first mutations in the COLIVA5 gene were reported in 1990 (Barker et al., 1990). Since that time, hundreds of

different mutations have been found in all parts of the COLIVA5 gene in affected families (Martin et al., 2000). These genetic abnormalities include deletions, insertions, and splice mutations. All of these mutations have been found to alter the COLIVA5 gene producing nonsense mutations that introduce premature stop codons or missense mutations that result in the substitution of amino acids. Frequently these missense mutations result in the substitution of a glycine residue. As discussed previously, glycine is necessary at every third residue for proper collagen folding and function. Major rearrangements in the COLIVA5 gene are very rare (Boye et al., 1991; Renieri et al., 1992). Even though linkage to the COLIVA5 gene has been established, mutations in the COLIVA5 gene can not always be found (Kawai et al., 1996).

Proper folding and subsequent trimer formation is essential for collagen type IV function. The loss of function is visualized in immunohistochemical studies that show the absence of staining of the product in target tissues. The collagen type IV molecules create distinct networks to form the structural framework of the basement membranes. The COLIVA5 product, which is usually present in the GBM, Bowman's capsule and distal tubule basement membrane, is absent in patients with X-linked AS (Nakanishi et al., 1994). COLIVA5 is also normally expressed in the lens capsule and Decemet's membrane of the eye, in the cochlea of the ear and in the skin. In patients with X-linked AS COLIVA5 is also reduced or absent in these tissues. Females who are heterozygous for COLIVA5 mutations frequently show abnormal staining of tissues expressing COLIVA5 (Kleppel et al., 1989a; Yoshioka et al., 1994) (Yoshioka et al., 1994). It is suggested that the variations in histochemical staining patterns in females heterozygous

for COLIVA5 mutations is due to differential lyonization in these tissues of the COLIVA5 locus. It is important to note that the absence of COLIVA5 staining in X-linked AS correlates with the absence of COLIVA3 and COLIVA4 staining in tissues where these genes are also normally expressed including the GBM (Nakanishi et al., 1994).

Mutations in the COLIVA3 and COLIVA4 genes are responsible for autosomal recessive AS. The first study to show aberrations in the COLIVA3 and COLIVA4 genes reported findings from four families (Mochizuki et al., 1994). Two of the families had mutations in the COLIVA3 gene and two had mutations in the COLIVA4 gene. Both of the mutations in the COLIVA3 gene were found in young females diagnosed with juvenile AS and each of these patients had GBM structures characteristic of AS. Both patients also had sensorineural loss of hearing. One of the COLIVA3 mutations found was a five base deletion that resulted in a premature stop codon in the NC-1 domain. The other COLIVA3 mutation was also a nonsense mutation in the NC-1 domain, but resulted from a single C to T transition. Mutations in the COLIVA4 gene were also identified from young female patients with juvenile AS from two other families, but unlike the positioning of the mutations in the NC-1 domain of the COLIVA3 genes above, both of the COLIVA4 mutations were found in areas coding for collagenous sequences. One COLIVA4 mutation was a G-A transition that changed a glycine to a serine and the other, found in a different region, was a single C-A transversion that resulted in a premature stop codon. Several other mutations have been reported in the COLIVA3 and COLIVA4 genes in patients diagnosed with autosomal AS (Boye et al.,

1998; Ding et al., 1995; Heidet et al., 2001; Lemmink et al., 1994; Longo et al., 2002). The above studies show that like the mutations in the COLIVA5 gene in X-linked AS described above, mutations in the COLIVA3 and COLIVA4 have been found in all parts of these genes and are not confined to any one area.

Thin basement membrane disease (TBMD) is also known as benign familial hamaturia and, as the name suggests, is characterized by a thinning of the GBM and is inherited as an autosomal dominant trait in humans (Gauthier et al., 1989; Rogers et al., 1973). TBMD has been linked to the COLIVA3/COLIVA4 locus (Buzza et al., 2001b; Lemmink et al., 1996). Twelve different mutations in the COLIVA4 gene in twelve different families with TBMD have been identified (Buzza et al., 2001a; Lemmink et al., 1996; Ozen et al., 2001)(Buzza 2003). Mutations in the COLIVA3 gene in families with TBMD have also been reported (Longo et al., 2002). Again, the mutations in these genes include nonsense mutations resulting from deletions and insertions, or missense mutations resulting in glycine substitutions that disrupt the triple helical nature of the molecule and the mutations are not localized and have been found throughout these genes. Mutations in either the COLIVA3 and COLIVA4 genes exist as heterozygotes in individuals with TBMD and mutations in both genes existing simultaneously have not been reported. Carriers of X-linked AS may also show thinning of the GBM (Rumpelt et al., 1974). The heterozygous nature of the mutations suggests that individuals diagnosed with TBMD may be carriers of autosomal recessive AS.

An autosomal dominant form of AS has been described in humans (Feingold et al., 1985). These families have a less severe phenotype with a slower rate of progression

as compared to patients with X-linked AS (Pochet et al., 1989). Analysis of a large family with autosomal dominant AS showed strong linkage with the COLIVA3/COLIVA4 locus, exclusion of the COLIVA5 gene and segregation of a COLIVA3 haplotype with the disease (Jefferson et al., 1997). In this family, a splice site mutation resulting in the skipping of exon 21 of the COLIVA3 gene has been reported (van der Loop et al., 2000). This mutation exists as a heterozygote with one normal allele present. The splice mutation of the affected gene does not alter the reading frame, but does result in a COLIVA3 chain with an internal deletion. Since the NC-1 domain is intact, this chain may be incorporated into the GBM and its presence may disrupt the intermolecular framework and interfere with the basement membrane architecture.

Goodpasture's syndrome (GS) is an autoimmune disorder that is characterized by circulating IgG antibodies targeting the GBM (Weber, 1993). Although the disease was first identified in 1919 by Dr. Earnest Goodpasture, it was not until 1987 that the target of the circulating antibodies was identified as a novel chain of type IV collagen (Butkowski et al., 1987). In fact, the colIVα3 molecule was first identified by co-immunoprecipitation with circulating GS antibodies and subsequent biochemical characterization (Saus et al., 1988). The specific epitope of the anit-GBM antibodies in GS has been identified as a cryptic epitope within the NC-1 region of ColIVα3 molecules (Netzer et al., 1999a). Although no outright cause of GP has been reported, it is possible that this epitope is revealed after exposure to GBM damaging agents including the inhalation of gasoline or acute infections.

Nail patella syndrome (NPS) is a pleiotropic disorder affecting nails, skeletal system, eyes and kidneys in humans (Beals and Eckhardt, 1969). NPS is also known as hereditary onychodysplasia (HOOD), Fong's disease, Turner-Kiesler syndrome or Osterreicher-Turner syndrome. In NPS, renal pathology involves defects in the GBM and the first sign of renal involvement in NPS is usually proteinuria with or without haematuria and is the main cause of death in these patients. (Gubler et al., 1980). GBM changes in NPS include irregular thickening with electron lucent loci resulting in a "moth eaten" appearance. The cause of NPS in humans is associated with defects in the LMX1B gene (Dreyer et al., 1998). The LMX1B gene has been identified as a transcription factor responsible for increased COLIVA3 and COLIVA4 expression that has binding activity in the first exon of the COLIVA4 gene, and LMX1B knockout mice show a decrease in COLIVA3 and COLIVA4 expression (Morello et al., 2001). Transcriptional induction of genes transcribed by LMX1B is required in podocyte differentiation as indicated by LMX1B knockout mice (Miner et al., 2002).

The above studies show that although the COLIVA3, COLIVA4 and COLIVA5 chains are referred to as the "minor" species of collagen, they have important structural roles in the GBM and that disruption of their function results in a wide range of pathologies.

Canine Hereditary Nephritis

Several models for HN have been described with specific indications of GBM involvement. The models that will be discussed have been reported as progressive kidney disorders with end stage renal disease (ESRD) and have other characteristics similar to human AS.

An example of HN has been described in kindred Samoyed dogs (Jansen et al., 1984). In this model, males spontaneously develop renal failure and their GBM architecture has extensive multilamellar splitting when examined by electron microscopy similar to humans diagnosed with AS (Jansen et al., 1986a; Jansen et al., 1986b). Affected male Samoyed dogs have glomerular changes that are associated with increased proteinuria that increases with the age of the dog and terminates in renal failure by 15 months of age. Carrier female Samoyed dogs were identified on the basis of having given birth to affected males. The carrier females do not regularly develop ESRD (although some cases have been reported after five years of age) and GBM structural analysis by electron microscopy show less extensive aberrations with focal lamellation which does not progress over time.

The rate of progression in the Samoyed model was studied using a selective breeding program (Jansen et al., 1986a). In early development, affected males and carrier females appeared healthy. Proteinuria was not detected until 2-3 months of age in either population, but all affected male dogs died by 15 months of age. The

predominance of affected males indicated an X-linked mode of inheritance which was confirmed with a later study (Jansen et al., 1986b).

An interesting study that would foretell many aspects of GBM physiology was conducted to determine the immunohistochemical nature of Samoyed HN (Thorner et al., 1987). This study was conducted without the knowledge of the existence of the minor species of collagen, but these investigators used Goodpasture syndrome (GS) antibodies, along with antibodies directed against laminin, collagen type IV and fibronectin. The only COLIV molecules known at that time were products of the COLIVA1 and COLIVA2 genes. The above antibodies were used in immunohistochemical studies with electron microscopy evaluation to characterize staining patterns in normal adult dogs, neonatal dogs, female carriers of Samoyed HN and male Samoyeds affected with HN. The results for laminin and COLIV antibodies showed strong staining of these products in the GBM in all of the populations above. Fibronectin staining in the GBM was positive for neonatal dogs and affected male Samoyeds, but was focal in carrier females and negative in normal dogs. This is an important finding in that with our current knowledge of GBM components, fibronectin is absent from normal GBM and agrin is the predominant HSPG in mature GBM. GS antibodies did not stain neonatal tissue, only stained in focal areas of carrier female GBM, was greatly reduced in affected males, but did stain normal dog GBM. We now know that the target of GS antibodies is an NC-1 epitope of the colIVα3 molecule. Also, it is now recognized that the colIVα3, colIVα4 and colIVα5 chains are negative for staining in patients with AS (Nakanishi et al., 1994). These and other studies indicated

that Samoyed HN was a model for X-linked AS in humans (Baumal et al., 1991; Thorner et al., 1989). After the molecular basis of X-linked AS was revealed to be caused by mutations of the COLIVA5 gene on the X chromosome (Barker et al., 1990), an investigation of this gene was directed against this gene in the Samoyed HN model (Zheng et al., 1994). This study was conducted using the known human COLIVA5 sequence and, using these sequences, primers were developed for the amplification, cloning, and sequencing of the canine COLIVA5 gene. The result was the identification of a single G-T transversion that resulted in a premature stop codon in exon 35 of the gene, thus confirming Samoyed HN as a model for human X-linked AS.

A second form of X-linked HN in a mixed breed of dogs has also been described (Lees et al., 1999). Because this kindred is of a mixed breed and originates from Navasota, Texas, this population has been termed NAV dogs. Individuals of this population were referred to the Texas A&M Veterinary Medical Center for evaluation because of kidney failure. The NAV model of HN has many features similar to the Samoyed model above. Pedigree analysis of the NAV dogs indicated an X-linked form of transmission. Male NAV dogs affected with HN exhibit proteinuria at three months of age and usually die before 15 months of age. Carrier females have proteinuria but do not usually progress to renal failure. Ultrastructurally, GBM aberrations of affected NAV dogs include thickening and multilamellar splitting. Since the NAV dogs were of mixed breed origin, migration of the Samoyed allele responsible for HN into the NAV population could not be excluded. The COLIVA5 locus responsible for HN in the Samoyed model was sequenced in affected males and determined not to be responsible

for HN in the NAV population. Thus, a different mutation responsible for HN in NAV dogs was suggested and established NAV HN as a separate model from Samoyed HN.

A form of HN in bull terriers has also been reported (Hood et al., 1995). Proteinuria is progressive and a high content of albumin in the urine strongly suggests involvement of the GBM. Glomerular changes in Bull terriers affected with HN have been reported. Chronic renal disease in this breed ranges from one to eight years of age (Robinson et al., 1989). Unlike other canine models, bull terrier HN is inherited as an autosomal dominant trait (Hood et al., 1990). Almost equal proportions of affected males and female bull terrier HN have been reported. Although autosomal dominant Alport syndrome in one human family has been determined to be caused by a splice site mutation in the COLIVA3 gene, immunohistochemical analysis of the GBM in bull terriers affected with HN show the presence of the COLIVα3 chain (Hood et al., 1995). A similar model for ADAS has been identified in a naturally occurring kindred in Dalmatians (Hood et al., 2002). Defects in the colIVα4 chain have been implicated in these models.

A form of NH has been described in English cocker spaniel dogs (ECS) that was initially termed familial nephropathy (Robinson et al., 1985; steward, 1984). These studies were initiated after presentation of clinical cases were referred to small animal practices and included both affected males and females. They described a rapid progression and proteinuria. Although no EM was attempted to characterize the GBM, the urine protein was determined to be of glomerular origin due to the presence of albumin and globulin (Steward, 1984). A study to characterize HN in ECS reported

findings by EM in affected dogs that indicated the extensive thickening and multilamellar splitting of the GBM characteristic of AS in humans (Lees et al., 1997). Although previous studies reported ultrastructural changes in the GBM associated with canine glomerulonephritis, these studies were not specific to HN in ECS (Macdougall et al., 1986; Murray and Wright, 1974). To study the early features and progression of ECS HN, known carriers were identified and two litters were produced by selective breeding (Lees et al., 1998b). The result of this breeding included one litter with a single male and one litter with one male and two females. Two of the dogs from this breeding (both females) developed HN , and the third dog (a single male) was normal at 30 months of age when this study was completed. After onset of proteinuria, loss of kidney function progressed rapidly and all of the affected dogs demonstrated renal lesions by electron microscopy.

Further investigation indicated ESC HN as a model for autosomal recessive AS (Lees et al., 1998b). This study included normal dogs, obligate carrier ECS dogs and affected ECS dogs. Proteinuria was first detected at five to six months of age in affected ECS dogs, however the rate of progression varied in this population. The development of azotemia in affected dogs varied from two to nine months and two dogs that were identified as affected had proteinuria that was significant, but transient. The pattern of immunohistochemical staining, however, was the same for every affected ECS dog. Laminin staining was universal in normal, ECS carriers and in affected ECS dogs. In affected ECS dogs, the GBM staining of the colIVα1 and colIVα2 chain was increased, while the colIVα5 staining was greatly decreased and the colIVα3 and colIVα4 chains

were totally absent. By contrast, staining of normal dog GBM showed a predominance of colIVα3, colIVα4 and colIVα5 chains and lesser staining for the colIVα1 and colIVα2 chains. Obligate carriers had a decreased colIVα5 staining. This is the same pattern of expression reported in the GBM of humans diagnosed with autosomal recessive AS (Gubler et al., 1995). Interestingly, the colIVα6 chain appeared to be faintly expressed in both normal and affected ECS dogs.

The detection of the colIVα6 chain in the canine GBM prompted further investigation (Lees et al., 2000). Immunolabeling using antibodies monospecific for the colIVα6 chain showed positive staining in the GBM of dogs less than 30 months of age, but positive in dogs more than 45 months of age. The population of dogs in this study were individuals from many breeds showing that the colIVα6 chain is not breed dependent. Human subjects were also included in this investigation, but regardless of age (30 years of age to 100 years of age), the colIVα6 chain was not detected. In humans an isoform switch that occurs during fetal development has been reported (Kalluri et al., 1997). In fetal GBM, the colIVα1 and colIVα2 chains are the only colIV isoforms present, but during development, the colIVα3, colIVα4 and colIVα5 chains are produced and, eventually, become the predominant isoform species in the mature GBM. It is suggested that the identification of the colVα6 chain in the GBM of older dogs represents a second isoform switch in this species. Because the colIVα6 chain was detected in the GBM of ECS dogs affected with HN less that 25 months of age, it is also

suggested that the isoform switch can occur at an earlier age in response to glomerular injury (Lees et al., 1998b).

Summary

Six different isoforms of Type IV collagen (colIVα1-6) have been identified. The individual isoforms of colIV are termed alpha chains and are translated from six different colIV genes (COLIVA1-A6). Collagen Type IV gene products compose the structural framework of basement membranes. The glomerular basement membrane (GBM) is a specialized basement membrane involved in the ultrafiltration processes of the kidney. The colIVα1-α5 chains are expressed in the human GBM while the colIV α1-α6 chains are expressed in the canine GBM. Many inherited diseases of the kidney have been reported and mutations in genes regulating kidney function have been identified. Alport syndrome (AS) is the most common form of human hereditary nephritis (HN). AS is defined as an inherited progressive kidney disorder associated with sensoneural deafness and is characterized by extensive thickening and multilamminar splitting of the GBM when examined by electron microscopy. AS has both X-linked (XLAS) and autosomal (ARAS) modes of inheritance. Mutations in the COLIVA5 gene are responsible for XLAS. A form of HN with characteristic splitting of the GBM with X-linked inheritance has been described in Samoyed dogs. A specific mutation in the COLIVA5 gene has been identified in Samoyed dogs affected with HN. Mutations in the

COLIVA3 and COLIVA4 genes are responsible for ARAS. A form of HN has been identified in English cocker spaniel dogs (ECS) that has been described as autosomal in inheritance and includes GBM abnormalities including extensive lammination characteristic of ARAS. The HN in ECS has been hypothesized to be a model for human ARAS. As such, we have isolated and characterized canine COLIVA3 and COLIVA4 sequences from normal dogs and ECS dogs affected with HN and compared the coding regions of these candidate genes.

CHAPTER II

THE ISOLATION AND CHARACTERIZATION OF CANINE COLIVA3

Introduction

Collagens are found in all vertebrate species as well as in Drosophila, sea urchins and nematodes (Blumberg et al., 1988). As molecules, the collagens are described as triple helical collagenous domains with the amino acid glycine at every third residue that are expressed extracellularly where they perform structural roles. To date, 19 different types of collagen have been identified (Uitto et al., 1986). Six different isoforms of type IV collagen (colIVα1-α6) have been described (Hudson et al., 1993). The individual isoforms of colIV are termed alpha chains and are translated from six different genes (COLIVA1-6).

Collagen type IV gene products compose the structural framework of basement membranes. Other basement membrane components include laminin, entactin/nidogen and heparine sulfate proteoglycan (HSPG). The glomerular basement membrane (GBM) is a specialized structure created predominately by epithelial podocytes in the kidney. The GBM is an important structure for the ultrafiltration processes of the kidney.

Hereditary nephritis (HN) refers to a group of inherited glomerular diseases of collagen type IV that often progress to renal failure. In humans, Alport syndrome (AS) is

the most common form of HN (Grunfeld, 1985). About 85% of AS families show X chromosome-linked inheritance (XLAS); the remainder show autosomal recessive (ARAS) or autosomal dominant inheritance (ADAS) (Reeders, 1992).

Several forms of canine HN have been described. HN in the Samoyed breed of dog has been described as a model of XLAS and a mutation in the COLIVA5 gene has been identified as causative for this disease (Zheng et al., 1994). In bull terriers, an autosomal dominant form of HN has been reported, although the molecular basis of disease has yet to be described (Hood et al., 1995).

A form of HN has also been described in English cocker spaniel (ECS) dogs (Lees et al., 1998a). ECS HN is characterized as a progressive kidney disease with proteinuria that often results in chronic renal failure. In ECS HN, the GBM has extensive thickening and splitting that increases with the progression of the disease. Humans with ARAS also have extensive thickening and splitting of the GBM. Mutations in the COL4A3 and COL4A4 genes have been identified in individuals with ARAS.

The similarities of the glomerular changes and the autosomal recessive mode of inheritance in both ECS HN and humans with ARAS suggest genetic similarities. As such, we have hypothesized that COL4A3 and COL4A4 are candidate genes in ECS HN. We have isolated and characterized the coding regions of these candidate genes in normal dogs (without HN) and compared these sequences from ECS dogs diagnosed with HN. This chapter describes the isolation and characterization of the COL4A3 coding regions in these populations.

Materials and Methods

Identification of Affected and Normal Populations. Each affected ECS dog met two criteria for the diagnosis of HN. First, the kidney disease exhibited clinicopathalogic features identical to those described previously as familial nephropathy in ECS. Second, glomerular examination with transmission electron microscopy (TEM) that exhibited extensive thickening and multilammellar splitting was determined to be pathogmonic of HN. The form of HN in ECS has been determined to be a breed specific disease (Lees et al., 1998b). Three individual dogs older than two years of age from non-ECS breeds were identified for the normal population for this study. The dogs in the normal population included a mixed breed female, a female bloodhound and a male Labrador retriever. ECS dogs are a rare breed in the United States with a registered population of ~5,000 individuals. Because of the rarity and structured breeding practices of the ECS population, it is unlikely that alleles causing ECS-HN would be present in mixed bred populations.

Tissue Isolation. Affected dogs and two normal dogs (a female bloodhound and mixed breed) were sacrificed and ~1g sections of kidneys containing the renal cortex were snap frozen in liquid nitrogen, transported on dry ice and stored at –80°C. Tissue from one normal dog was from a testicle of a Labrador retriever obtained in the College of Veterinary Medicine Texas A&M University from a standard neutering procedure.

RNA Preparation. Total RNA was isolated from frozen tissue using Trizol (Sigma, St Louis, MO) according to manufacturers protocol. Briefly, ~1g of snap frozen tissue was ground in a metal beaker embedded in dry ice and ~100mg was added to 1ml of Trizol in a pyrex test tube. A plastic coated dounce homogenizer was used to macerate the tissue in the tube. The dounce was manipulated manually for approximately one minute at room temperature. To prevent shearing of the RNA, mechanical methods of dounce manipulation were avoided. The Trizol/tissue mixture was transferred to a 1.8ml microcentrifuge tube, allowed to rest at room temperature for ~5 minutes and centrifuged at 1,000xg for one minute to pellet any unsolubilized material. Occasionally this mixture was stored at –80°C, but immediate use is recommended.

To a new 1.8 ml microcentrifuge tube, .5 ml of the Trizol/tissue mixture and 100ul of cloroform was combined. The tube was briefly vortexed (< 5 seconds) and then centrifuged at 10,000xg for 10 minutes. 400ul of the clear upper liquid was transferred to a new 1.8ml microcentrufuge tube and 200ul of isopropanol was added, mixed briefly and centrifuged at 10,000xg for 10 minutes. The isopropanol was removed and the pellet was washed with 1ml of 75% ethanol. After the removal of ethanol, the pellet was resuspended in 50ul of RNAse free water and 3ul of RNAse inhibitor (Promega, Madison, WI) was added.

A portion of the RNA was run on a 1% agarose gel in TAE buffer and visualized by staining with ethidium bromide for qualitative analysis. For quantitative purposes, a portion of the RNA was also used for spectroscopic analysis at 280A° and 260A°.

cDNA Synthesis. The synthesis of cDNA was performed using the Superscript II first strand synthesis system (Invitrogen, Carlsbad, California) according to manufacturers suggestions. Briefly, 5 ug of total RNA was incubated with an oligo dT primer and 1uL of 10mM dNTP's at 70°C for 5 minutes, chilled on iced, and then combined with 4 uL 5X first strand buffer, 2 uL .1M DTT and10 units of Superscript reverse transcriptase at 45°C for one hour. The resulting product was used directly in PCR amplification reactions.

PCR Amplification. To assure the accuracy of the amplification process, the enzyme used for PCR amplification of the cDNA was AccuTaq LA DNA Polymerase (Sigma, St. Louis, MO) which has 3' proofreading ability. According to manufacturers suggestions, the amplification reaction contained 5 uL 10X PCR buffer, 2.5 uL 10mM dNTP mix, 600 nmol of each primer, .1 units of AccuTaq LA DNA Polymerase in a 50 uL reaction. As a template, 2 uL of the cDNA was also included. The amplification conditions for all reactions were an initial denaturation at 98°C for 30 seconds followed by repeated denaturation at 94°C for 30 seconds, annealing at 65°C for 30 seconds and a 5 minute extension at 68°C for 35 cycles. A 7 ul sample of the final reaction was analyzed on a 1% agarose gel using TAE buffer to check for the quality of the reaction.

Amplification Primers. Primers were designed from a human COLIVA3 sequence (Genebank Accession: NM000091). Primers for the targeted areas were analyzed using OLIGO 4.0 software (National Biosciences, Plymouth, MN) for dimerization, the

presence of hairpin loops and internal stability. Three sets of primers were used to amplify the canine COLIVA3 cDNA. The NC-1 region was amplified using COLIVA3 NC-1 Forward 5'CTG GAG AAA AAG GAC GTG ATG3' and COLIVA3 NC-1 Reverse 5' CAG CTT TCA CAG TTG ATG GAA TAG5' primers. The 5' region (to ~ 260 bases of the start codon) was amplified using COLIVA3 Forward 5'CCG ACT TGT CGG TGT ACC AGG ATG3' and COLIVA3 Reverse 5'CTC TCC TTC TTC CCC CTT GGC TCC3' primers. The intervening sequence was amplified using the COLIVA3 Forward 5'GTG ATT GGG ATG ATG GGC TAT C3' and COLIVA3 Reverse 5'TCT CCA ATT GGC CCA GGA GTT C3' primers.

Gel Purification. Amplified products were run on a 1% agarose gel and desired fragments were visualized with low intensity ultra violet light. Desired products were cut out of the gel using a clean razor blade and placed into a 1.8 ml microcentrifuge tube. The PCR product was purified using the QIAquick gel purification system (Qiagen, Valencia, CA) and eluted with 50 uL water.

Cloning. The gel purified amplification products were concentrated in a speed vacuum under centrifugation until the 50 uL of eluted product was reduced to approximatetly 5uL. The gel purified products were cloned into pT7 cloning vectors using perfectly blunt cloning kits (Novagen, Madison, WI) using T4 ligase. The cloned products were transformed into chemically competent cells and plated on LB media containing 50 ug/ ml ampicillin, 20 ug/ml tetracycline and 20 ug/ml IPTG that had been overlayed with 20

uL of 50 ug/ml X-gal. The presence of gene fragments in subsequent plasmids was confirmed by the presence of white colonies and restriction digestion using EcoRI and HindIII.

Sequencing. Sequencing reactions contained 4 uL Big Dye V3.0 (Applied Biosystems, Foster City, CA), 400 ng plasmid DNA and 50 nM primer in a 10 uL total reaction volume. The parameters for the sequencing reactions included incubation at 95°C for 10 seconds, 50°C for 5 seconds and 60°C for 4 minutes in a thermocycler for 44 cycles. The resulting reaction was purified on BioMax SPIN-50 mini-columns (BioMax, Odenton, MD). Briefly, the columns were placed in a microcentrifuge at 1000xg for 3 minutes. The columns were transferred to a new 1.8 ml microcentrifuge tube and the sequencing reaction placed on the column. The BioMax spin columns were again centrifuged at 1000xg for 3 minutes. The columns were discarded and the recovered sample was dried in a speed vacuum for 40 minutes at room temperature under darkened conditions. The sequencing reactions were then analyzed in a commercial sequencing facility.

Sequence Analysis. Chromatograms of the resulting sequence were collected and the sequences were downloaded into MacVector V4.0 software (Eastman Kodak Company, Rochester, NY). The sequences were alligned with the human COLIVA3 sequence (NM000091) to determine its relative position with respect to the human isoform. Canine sequences that were confirmed to be of COLIVA3 origin were transferred into AssemblyLine V1.0 software files (Accelrys, Princeton, NJ) for sequence alignment.

Consensus sequences were then transferred to SDSC Biology Workbench (workbench.sdsc.edu) for subsequent analysis.

Results

Sequences representing the coding regions of the normal canine COLIVA3 gene have been isolated by PCR amplification from cDNA derived from total RNA (Appendix 1). A total of 4,643 bases have been characterized by sequence analysis representing 92.6% of the human coding sequence. The complete COL4A3 cDNA from the human is 8,050 bases and has a coding region of 5,010 bases (Genbank accession:NM000091). The mouse COL4A3 cDNA has also been characterized and contains 7,765 bases with a 5,010 base coding region (Genbank accession: NM007734). While the remaining ~275 nucleotides of the 5' end of the canine COL4A3 coding region have yet to be characterized, the 4,643 bases represents a majority of this sequence. Nucleic acid alignment using LALINE sequence analysis (workbench.sdsc.edu) show 83.4% homology between the canine COL4A3 and humanCOL4A3 sequences and 80.0% homology between the canine COL4A3 and mouse COL4A3 sequences. The human COL4A3 and mouse COL4A3 sequences share 81.4% homology (Appendix 2).

The canine COL4A3 sequences that have been isolated code for 1,547 amino acids. The human COL4A3 gene codes for 1,671 amino acids and the mouse COL4A3 gene codes for 1,670 amino acids. Using LALINE sequence analysis

(workbench.sdsc.edu), the amino acid sequence homology between the normal canine COL4A3 sequences and the human COL4A3 sequences is 83.4% and 79.8% homologous to the mouse COL4A3 sequences (Appendix 3). The translated coding regions of the human and mouse COL4A3 gene have 78.8% homology.

Interruptions in the collagenous region are characteristic of type IV collagen molecules. The human has 23 interruptions and the mouse has 21 interruptions in this region. Twenty interruptions have been identified in the canine COL4A3 sequences in the collagenous domain. The presence of glycine at every third residue defines the collagenous sequence. Glycine represents 27.3% of the human COL4A3 coding sequences and 27.6% of the mouse COL4A3 coding sequences. The canine COL4A3 coding region contains 27.2% glycine. Proline is another common amino acid in collagenous seuences. Proline represents 19.2% of the human COL4A3 sequences, 20.6% of the mouse COL4A3 sequences and 20.5% of the canine COL4A3 sequences. The codons for glycine are very rich in guanine (GGG, GGA,GGC,GGT) and the codons for proline are very rich in cystosine (CCC, CCT, CCG, CCA). Thus, collagenous sequences are very GC rich. The canine COL4A3 coding region contains 57.1% GC while the human and mouse COL4A3 sequences each contain 56.3% and 57.2% GC respectively.

COL4A3 sequences have also been isolated from ECS dogs diagnosed with HN from cDNA derived from total RNA isolated from fresh frozen kidney. A 4,645 base COL4A3 consensus sequence has been characterized in this affected population representing 1,547 amino acids. The COL4A3 sequences from this population are almost

identical at the nucleotide level. The specific nature of these sequences will be the focus of a later chapter.

CHAPTER III

THE ISOLATION AND CHARACTERIZATION OF CANINE COLIVA4

Introduction

Collagens are found in all vertebrate species as well as in Drosophila, sea urchins and nematodes (Blumberg et al., 1988). As molecules, the collagens are described as triple helical collagenous domains with the amino acid glycine at every third residue that are expressed extracellularly where they perform structural roles. To date, 19 different types of collagen have been identified (Uitto et al., 1986). Six different isoforms of type IV collagen (colIVα1-α6) have been described (Hudson et al., 1993). The individual isoforms of colIV are termed alpha chains and are translated from six different genes (COLIVA1-6).

Collagen type IV gene products compose the structural framework of basement membranes. Other basement membrane components include laminin, entactin/nidogen and heparine sulfate proteoglycan (HSPG). The glomerular basement membrane (GBM) is a specialized structure created predominately by epithelial podocytes in the kidney. The GBM is an important structure for the ultrafiltration processes of the kidney.

Hereditary nephritis (HN) refers to a group of inherited glomerular diseases of collagen type IV that often progress to renal failure. In humans, Alport syndrome (AS) is

the most common form of HN (Grunfeld, 1985). About 85% of AS families show X chromosome-linked inheritance (XLAS); the remainder show autosomal recessive (ARAS) or autosomal dominant inheritance (ADAS) (Reeders, 1992).

Several forms of canine HN have been described. HN in the Samoyed breed of dog has been described as a model of XLAS and a mutation in the COLIVA5 gene has been identified as causative for this disease (Zheng et al., 1994). In bull terriers, an autosomal dominant form of HN has been reported, although the molecular basis of disease has yet to be described (Hood et al., 1995).

A form of HN has also been described in English cocker spaniel (ECS) dogs (Lees et al., 1998a). ECS HN is characterized as a progressive kidney disease with proteinuria that often results in chronic renal failure. In ECS HN, the GBM has extensive thickening and splitting that increases with the progression of the disease. Humans with ARAS also have extensive thickening and splitting of the GBM. Mutations in the COL4A3 and COL4A4 genes have been identified in individuals with ARAS.

The similarities of the glomerular changes and the autosomal recessive mode of inheritance in both ECS HN and humans with ARAS suggest genetic similarities. As such, we have hypothesized that COL4A3 and COL4A4 are candidate genes in ECS HN. We have isolated and characterized the coding regions of these candidate genes in normal dogs (without HN) and compared these sequences from ECS dogs diagnosed with HN. This chapter describes the isolation and characterization of the COL4A4 coding regions in these populations.

Materials and Methods

Identification of Affected and Normal Populations. Each affected ECS dog met two criteria for the diagnosis of HN. First, the kidney disease exhibited clinicopathalogic features identical to those described previously as familial nephropathy in ECS. Second, glomerular examination with transmission electron microscopy (TEM) that exhibited extensive thickening and multilammellar splitting was determined to be pathogmonic of HN. The form of HN in ECS has been determined to be a breed specific disease (Lees et al., 1998b). Three individual dogs older than two years of age from non-ECS breeds were identified for the normal population for this study. The dogs in the normal population included a mixed breed female, a female bloodhound and a male Labrador retriever. ECS dogs are a rare breed in the United States with a registered population of ~5,000 individuals. Because of the rarity and structured breeding practices of the ECS population, it is unlikely that alleles causing ECS-HN would be present in mixed bred populations.

Tissue Isolation. Affected dogs and two normal dogs (a female bloodhound and mixed breed) were sacrificed and ~1g sections of kidneys containing the renal cortex were snap frozen in liquid nitrogen, transported on dry ice and stored at –80°C. Tissue from one normal dog was from a testicle of a Labrador retriever obtained in the College of Veterinary Medicine at Texas A&M University from a standard neutering procedure.

RNA Preparation. Total RNA was isolated from frozen tissue using Trizol (Sigma, St Louis, MO) according to manufacturers protocol. Briefly, ~1g of snap frozen tissue was ground in a metal beaker embedded in dry ice and ~100mg was added to 1ml of Trizol in a pyrex test tube. A plastic coated dounce homogenizer was used to macerate the tissue in the tube. The dounce was manipulated manually for approximately one minute at room temperature. To prevent shearing of the RNA, mechanical methods of dounce manipulation were avoided. The Trizol/tissue mixture was transferred to a 1.8ml microcentrifuge tube, allowed to rest at room temperature for ~5 minutes and centrifuged at 1,000xg for one minute to pellet any unsolubilized material. Occasionally this mixture was stored at –80°C, but immediate use is recommended.

To a new 1.8 ml microcentrifuge tube, .5 ml of the Trizol/tissue mixture and 100ul of cloroform was combined. The tube was briefly vortexed (< 5 seconds) and then centrifuged at 10,000xg for 10 minutes. 400ul of the clear upper liquid was transferred to a new 1.8ml microcentrufuge tube and 200ul of isopropanol was added, mixed briefly and centrifuged at 10,000xg for 10 minutes. The isopropanol was removed and the pellet was washed with 1ml of 75% ethanol. After the removal of ethanol, the pellet was resuspended in 50ul of RNAse free water and 3ul of RNAse inhibitor (Promega, Madison, WI) was added.

A portion of the RNA was run on a 1% agarose gel in TAE buffer and visualized by staining with ethidium bromide for qualitative analysis. For quantitative purposes, a portion of the RNA was also used for spectroscopic analysis at 280A° and 260A°.

cDNA Synthesis. The synthesis of cDNA was performed using the Superscript II first strand synthesis system (Invitrogen, Carlsbad, California) according to manufacturers suggestions. Briefly, 5 ug of total RNA was incubated with an oligo dT primer and 1uL of 10mM dNTP's at 70°C for 5 minutes, chilled on iced, and then combined with 4 uL 5X first strand buffer, 2 uL .1M DTT and10 units of Superscript reverse transcriptase at 45°C for one hour. The resulting product was used directly in PCR amplification reactions.

PCR Amplification. To assure the accuracy of the amplification process, the enzyme used for PCR amplification of the cDNA was AccuTaq LA DNA Polymerase (Sigma, St. Louis, MO) which has 3' proofreading ability. According to manufacturers suggestions, the amplification reaction contained 5 uL 10X PCR buffer, 2.5 uL 10mM dNTP mix, 600 nmol of each primer, .1 units of AccuTaq LA DNA Polymerase in a 50 uL reaction. As a template, 2 uL of the cDNA was also included. The amplification conditions for all reactions were an initial denaturation at 98° Cfor 30 seconds followed by repeated denaturation at 94°C for 30 seconds, annealing at 65°C for 30 seconds and a 5 minute extension at 68°C for 35 cycles. A 7 ul sample of the final reaction was analyzed on a 1% agarose gel using TAE buffer to check for the quality of the reaction.

Amplification Primers. Primers were designed from a human COLIVA4 sequence (Genebank Accession: NM000092). Primers for the targeted areas were analyzed using OLIGO 4.0 software (National Biosciences, Plymouth, MN) for dimerization, the

presence of hairpin loops and internal stability. Three sets of primers were used to amplify the canine COLIVA4 cDNA. The NC-1 region was amplified using COLIVA4 NC-1 Forward 5'TTG GCC CTG GAT ACC TCA GTG3' and COLIVA4 NC-1 Reverse 5'GGC ACC TGC TGA TTT TCT GG3' primers. The 5' region (to -140 bases of the start codon in the human) was amplified using COLIVA4 Forward 5'TGA CCC AGA ACA CAG AAC CTC AC3' and COLIVA4 Reverse 5'CCC TTC TCT CCT TTT GGT CCA ATT C3' primers. The intervening sequence was amplified using the COLIVA4 Forward 5'AGA GGC GCT CCA GGG ATA C3' and COLIVA4 Reverse 5'CCT TCT GGT CCT CTT G3' primers for the 2,155 base region adjacent to the NC-1. The 1,590 base region adjacent to the 5' sequence was amplified with the COL4A4 F 5'GGT CTC CGG GTT CCT ATG3' AND COL4A4 R 5'CTG GCG GCC CAA GAA G3' primers.

Gel Purification. Amplified products were run on a 1% agarose gel and desired fragments were visualized with low intensity ultra violet light. Desired products were cut out of the gel using a clean razor blade and placed into a 1.8 ml microcentrifuge tube. The PCR product was purified using the QIAquick gel purification system (Qiagen, Valencia, CA) and eluted with 50 uL water.

Cloning. The gel purified amplification products were concentrated in a speed vacuum under centrifugation until the 50 uL of eluted product was reduced to approximatetly 5 uL. The gel purified products were cloned into pT7 cloning vectors using perfectly blunt cloning kits (Novagen, Madison, WI) using T4 ligase. The cloned products were

transformed into chemically competent cells and plated on LB media containing 50 ug/ ml ampicillin, 20 ug/ml tetracycline and 20 ug/ml IPTG that had been overlayed with 20 uL of 50 ug/ml X-gal. The presence of gene fragments in subsequent plasmids was confirmed by the presence of white colonies and restriction digestion using EcoRI and HindIII.

Sequencing. Sequencing reactions contained 4 uL Big Dye V3.0 (Applied Biosystems, Foster City, CA), 400 ng plasmid DNA and 50 nM primer in a 10 uL total reaction volume. The parameters for the sequencing reactions included incubation at 95°C for 10 seconds, 50°C for 5 seconds and 60°C for 4 minutes in a thermocycler for 44 cycles. The resulting reaction was purified on BioMax SPIN-50 mini-columns (BioMax, Odenton, MD). Briefly, the columns were placed in a microcentrifuge at 1,000xg for 3 minutes. The columns were transferred to a new 1.8 ml microcentrifuge tube and the sequencing reaction placed on the column. The BioMax spin columns were again centrifuged at 1,000xg for 3 minutes. The columns were discarded and the recovered sample was dried in a speed vacuum for 40 minutes at room temperature under darkened conditions. The sequencing reactions were then analyzed in a commercial sequencing facility.

Sequence Analysis. Chromatograms of the resulting sequence were collected and the sequences were downloaded into MacVector V4.0 software (Eastman Kodak Company, Rochester, NY). The sequences were alligned with the human COLIVA4 sequence (NM000092) to determine its relative position with respect to the human isoform.

Canine sequences that were confirmed to be of COLIVA3 origin were transferred into AssemblyLine V1.0 software files (Accelrys, Princeton, NJ) for sequence alignment. Consensus sequences were then transferred to SDSC Biology Workbench (workbench.sdsc.edu) for subsequent analysis.

Results

Sequences representing the coding regions of the normal canine COLIVA4 gene have been isolated by PCR amplification from cDNA derived from total RNA (Appendix 5). A total of 5,163 bases have been characterized by sequence analysis representing ~98% of the human coding sequence. The complete COL4A4 cDNA from the human is 5,726 bases and has a coding region of 5,073 bases (Genbank accession:NM000092). The mouse COL4A4 cDNA has also been characterized and contains 5,918 bases with a coding region of 5,049 bases (Genbank accession: AF169338). While the remaining ~50 nucleotides of the 3' end of the canine COL4A4 coding region have yet to be characterized, the 5,163 bases represents a majority of this sequence. Nucleic acid alignment using LALINE sequence analysis (workbench.sdsc.edu) show 85.0% homology between the canine COL4A4 and human COL4A4 sequences and 79.0% homology between the canine COL4A4 and mouse COL4A4 sequences. The human COL4A4 and mouse COL4A4 sequences share 80.7% homology (Appendix 6).

The canine COL4A4 sequences that have been isolated code for 1,677 amino acids. The human COL4A4 gene codes for 1,691 amino acids and the mouse COL4A4 gene codes for 1,683 amino acids. Using LALINE sequence analysis (workbench.sdsc.edu), the amino acid sequence homology between the normal canine COL4A4 sequences and the human COL4A4 sequences is 83.2% and 78.3% homologous to the mouse COL4A4 sequences (Appendix 7). The translated coding regions of the human and mouse COL4A4 gene have 78.6% homology.

Interruptions in the collagenous region are characteristic of type IV collagen molecules. The human has 24 interruptions and the mouse has 23 interruptions in this region. In the canine COL4A4 sequences, 23 interruptions have been identified in the collagenous domain. The presence of glycine at every third residue defines the collagenous sequence. Glycine represents 28.0% of the human COL4A4 coding sequences and 27.5% of the mouse COL4A4 coding sequences. The canine COL4A4 coding region contains 27.8% glycine. Proline is another common amino acid in collagenous seuences. Proline represents 20.8% of the human COL4A4 sequences, 20.9% of the mouse COL4A4 sequences and 21.2% of the canine COL4A4 sequences. The codons for glycine are very rich in guanine (GGG, GGA,GGC,GGT) and the codons for proline are very rich in cystosine (CCC, CCT, CCG, CCA). Thus, collagenous sequences are very GC rich. The canine COL4A4 coding region contains 60.5% GC while the human and mouse COL4A4 sequences each contain 59.3% and 59.7% GC respectively.

COL4A4 sequences have also been isolated from ECS dogs diagnosed with HN from cDNA derived from total RNA isolated from fresh frozen kidney. A 5,011 base COL4A4 consensus sequence has been characterized in this affected population representing 1,607 amino acids. The COL4A4 sequences from this population are almost identical at the nucleotide level. The specific nature of these sequences will be the focus of a later chapter.

CHAPTER II

SUMMARY AND CONCLUSIONS

The problematic nature of the identification of mutations in ECS HN can be derived from studies that have attempted to find mutations in the COL4A3 and COL4A4 genes in ARAS. While the detection rate of mutations in the COL4A5 coding region in XLAS has reached ~90%, it has been reported that perhaps 10% of the mutations are in non-coding regions of the CL4A5 gene (Martin et al., 1998). The 90% detection rate of mutations in the COL4A5 gene in XLAS does not represent the current rate of detection of COL4A3 and COL4A4 mutations in ARAS. In the investigation of the COL4A3 or COL4A4 genes in ARAS, the success rate of finding mutations in these genes individually is very low. When both the COL4A3 and COL4A4 genes have been analyzed together in ARAS, the identification of causal mutations is increased, but does not reach 100% detection in these genes. Together, two studies that analyzed both the COL4A3 and COL4A4 genes in ARAS have shown less than a 30% rate of detection of mutations in these genes. This low rate of detection may represent the limitations in the methodology of exon screening which, when applied to the COL4A5 gene in XLAS, has an equally low rate of detection.

In both studies, diagnosis of AS had been made including the splitting of the GBM, and exclusion of the COL4A5 gene had been determined. Although selective

criteria were used to diagnose ARAS in these studies, this methodology represents a more permissive basis of diagnosis as compared to the strict interpretation of ECS-HN. In one study (Mochizuki et al., 1994), mutations in the COL4A3 and COL4A4 genes in 4 of 7 families was reported. These included two mutations in the COL4A3 gene (both nonsense mutations) and two mutations in the COL4A4 gene (one glycine substitution and one nonsense mutation). In another study (Longo et al., 2002), 8 mutations were identified in 30 probands. This study included six probands with ADAS presumed to be of COL4A3 or COL4A4 origin. While Mochizuki et al. reported an equal number of COL4A3 and COL4A4 mutations, Longo et al. reported seven COL4A3 mutations (two nonsense mutations, three microdeletions a glycine substitution, and a missense at the start codon) and only one COL4A4 mutation (one glycine substitution). The combined detection rate in these studies for mutations in either the COL4A3 or COL4A4 genes is 12/37. This low rate of detection suggests that either mutations in the COL4A3 and COL4A4 genes in ARAS are difficult to identify or, possibly, the presence of another genetic locus responsible for this pathology.

Indication of the latter possibility is represented in a study of haplotype segregation in thin basement membrane disease (TBMD) (Buzza et al., 2001b). While it is possible to perform linkage analysis of haplotype segregation of the COL4A3/COL4A4 locus, it is difficult to differentiate segregation patterns to any one of these loci. In both man and mouse, the COL4A3 and COL4A4 genes lie head to head where they share a common promoter and therefore can be considered to be examples of very close linkage. Haplotype segregation of genomic microsatellites of the

COL4A3/COL4A4 locus and the COL4A5 locus were studied in the families of 22 individuals diagnosed with TBMD. In this study, 35% of the families showed linkage to the COL4A3/COL4A4 locus, 18% showed linkage to the COL4A5 locus, but 45% of the families failed to show linkage to either locus. TBMD is often inherited as an autosomal dominant trait and, as the name indicates, is characterized by a thinned GBM (Blumenthal et al., 1988). Heterozygous carriers for causal mutations in ARAS (Lemmink et al., 1996) and female carriers of XLAS (Rumpelt et al., 1974) have clinocopathalogic symptoms identical to those of TBMD including a thinning of the GBM. It is suggested that TBMD represents a carrier state for ARAS or XLAS (Netzer et al., 1999b). Incomplete penetrance or improper diagnosis may have affected the lack of segregation of the either the COL4A3, COL4A4 or COL4A5 genes with TBMD. These results do not preclude, however, the presence of another locus responsible for TBMD: a recognized disease of the GBM. The possibility of another genetic locus responsible for ARAS might be resolved with similar haplotype analysis to determine the frequency of segregation of the COL4A3/COL4A4 locus in these families. Similarly, haplotype analysis could help to determine COL4A3/COL4A4 segregation in ECS HN.

Implication of either the COL4A3 or COL4A4 gene in ECS HN by analysis of haplotype segregation in our current study is very difficult. The experimental design in this study was to determines sequences in alleles not responsible for ECS HN and compare these sequences in ECS-HN populations. Optimally, haplotype analysis would be conducted in individuals of the same breed. This is possible in the analysis of the X-linked HN in the Samoyed breed. In this case, it is possible to positively identify affected

males, normal males and carrier females. Similar haplotype analysis is complicated in our study because of the presumed autosomal nature of ECS HN. It is presumed that affected individuals are homozygous recessive in ECS HN populations. Because the total ECS population in the United States is less than 5,000 individuals, there exists a possibility that any one individual may be a heterozygous carrier of ECS HN alleles. ECS HN has been reported to be breed specific. No other autosomal recessive HN has been reported in any other canine populations. For these reasons, non-ECS populations were selected for the absence of ECS HN alleles. Therefore, haplotype analysis in this study may represent breed specific ECS haplotypes and not represent markers specific for ECS HN. Segregation of haplotypes is often used for the analysis of quantitative trait loci (QTL). The experimental design for these studies involving haplotype segregation is very different from the approach of our study. Haplotype analysis of QTL does not frequently use genetic elements derived from cDNA that has been cloned.

While the vast majority of the COL4A3 and COL4A4 coding regions have been identified, it is possible that mutations exist in other areas yet to be characterized. These include ~250 bases at the 5' end and ~30 bases at the 3' end of the COL4A3 coding region, and ~50 bases at the 5' end and ~50 bases at the 3' end of the COL4A4 coding region. Other regions of interest include promoter elements, 5' and 3' UTRs, exon/intron junctions and intronic sequences.

Several differences have been identified between the normal and ECS-HN COL4A3 sequences. The effects of these individual nucleotide changes with regards to codon changes are summarized in Appendix 9. Several differences have also been

identified between the normal and ECS-HN COL4A4 sequences. The effects of these individual nucleotide changes with regards to codon changes are summarized in Appendix 10.

The most significant differences between normal canine populations and ECS-HN populations in the COL4A3 and COL4A4 sequences are two base changes in the COL4A3 allele. One difference that has been identified is a G-A transition that results in an amino acid change at position ~3,250 (See Appendix 7). While the codon for the normal sequence is AGG and codes for arginine, the G-A transition results in an AAG codon that codes for lysine. The sequence difference has been identified in every affected ECS examined (N=3) from 4 different clones and was present in each of 14 sequence reactions. This transition was identified in two normal non-ECS individuals from 4 clones in 7 sequencing reactions. Another significant difference identified in the COL4A3 sequences is another single base change at position ~3,220 (See Appendix 7). While the normal sequence is AGC and codes for serine, a G-T transversion has been identified that changes the codon in ECS alleles to ATC to code for isoleucine. This transversion was identified in three ECS-HN individuals from 4 different clones and was present in each of 12 sequencing reactions. The normal sequence has been confirmed in two normal individuals from four clones in 6 sequencing reactions. It is difficult to determine if these differences are causal mutations for ECS HN. If identification of heterozygosity in known carriers in these regions could be determined between normal and affected populations, these markers could be very informative. Based on these sequences, primers have been designed from regions flanking areas of interest to amplify

these regions from genomic DNA (gDNA) isolated from circulating lymphocytes for the identification of heterozygous alleles at this locus. The process has been difficult because the primary gDNA sequence of these genes has not been determined. Based on corresponding human sequences, both the COL4A3 and COL4A4 genes should each cover over 150 kb each and contain approximately 50 introns. Because the position and size of introns and exons are unknown for the canine genes, it is difficult to design parameters for amplification reactions. Although this does not preclude the possibility of isolating these sequences from gDNA, the process has proven difficult. An alternative approach to identify heterozygosity in these loci from known carrier populations is to isolate these sequences from the coding regions of these genes. This could include the amplification, cloning and sequencing of these regions from cDNA derived from total RNA. Because of the difficulties involved in the specific amplification of collagen genes, direct sequencing of amplification products with the primers that have been designed has proven difficult. Once the primary sequence of the gDNA has been determined, more specific primers for these areas of interest can be synthesized for direct sequencing of gDNA amplification reactions. Since both alleles may be equally expressed, it is possible to identify the two alleles by sequencing cloned sequences from products amplified from cDNA. Using the binomial equation ([N!/X!(N-X)!](Px)∧x(PN-x)∧N-x), greater than 95% detection of heterozygous loci can be determined when 5 clones are analyzed from products of individual amplification reactions.

The question as to the origin of the ECS-HN allele is an interesting topic for discussion. Clinically, ECS-HN is not indicated by spontaneous mutation, but rather by

an established genetic aberration in the germ line of this population. ECS-HN has been reported to date to be a breed specific pathology indicating that the ECS-HN allele is not present in other purebred populations and, within the ECS breed, is restricted to “party colored” individuals. Purebred canine populations exist as groups indicating a common origin. Examples of purebred groups include hounds, terriers and retrievers among others. Examples of purebred spaniel populations include the ECS as well as the American cocker spaniel, the Brittany spaniel, and Springer spaniel. No indication of the ECS-HN allele, however, is present in these populations. It is difficult to determine whether the ECS-HN allele was present during the establishment of the ECS population or originated during ECS breed development. If the ECS-HN allele was present during the establishment of the population, then this causal mutation might have also been present in other populations. If the ECS-HN allele was present in other populations, especially other spaniel populations, many factors could have caused its disappearance in non-ECS breeds. Selective forces could have eliminated the allele in non-ECS populations since the ECS-HN allele causes a decreased fitness. However, the perpetual nature of the ECS-HN allele in ECS populations could be due to the recessive nature of the allele that would mask its presence in heterozygous populations making selection difficult. Because these populations are small and because inbreeding tends toward genetic homozygosity, the ECS-HN allele could also have been lost by genetic drift. Alternatively, the ECS-HN allele could have been introduced in the evolution of this population.

Recessive alleles could be perpetuated through breeding practices inherent in the establishment and growth of purebred canine populations. Most purebred canine populations arise from a small number of individuals. Often there is a major sire that may be bred to many females. This practice can lead to a high frequency of alleles from a small portion of the breeding population. In this case, the frequency of the ECS-HN allele may have been amplified in this population by the founder effect. This would suggest that the ECS-HN allele, by whatever means of introduction, would be identical by descent. In pure genetic terms, identity by decent infers loss of all genetic heterogeneity to create identical alleles. This does not necessarily have to be a strict definition. Identity by decent could also infer homozygosity at one locus causing a particular phenotype, but genetic polymorphism at other allelic loci arising from subsequent mutations. The COL4A3 and COL4A4 sequences we have characterized from ECS-HN populations are very homozygous as are the sequences generated from non-ECS populations. Decreased homozygosity is expected in small-inbred populations. The polymorphisms that have been identified in the ECS-HN populations do not preclude, however, homozygosity at the ECS-HN locus. Although the ECS population has been recently established (less than 500 years ago), many hundreds of generations may have occurred thus creating the opportunity for the introduction of polymorphisms. The theory of the molecular clock suggests a constant rate of molecular evolution. Although studies on the molecular clock frequently focus on populations separated genetically over extreme time periods, it is not inconceivable that a causal mutations for ECS-HN may have been introduced into the population’s, albeit, brief history.

It is difficult to establish the transitional change that substitutes an arginine for a lysine as the causal mutation in ECS-HN. This change does not result in a nonsense mutation, does not occur at an essential structural element such as glycine and does not drastically effect a proline residue adjacent to a glycine or substitute a non-conserved amino acid. The neutral theory of molecular evolution, first established by Kimura, suggests that most mutations are neutral. The neutrality theory is indicated at two levels. First, that genetic mutations often do not change the amino acid (due to the wobble effect), and, second, when mutations do change the coding sequence they often substitute neutral changes that do not effect protein structure or function. The theory is based on the three types of mutations, namely, deleterious mutations, neutral mutations and advantageous mutations. Deleterious mutations are frequently eliminated from populations, advantageous mutations are exceedingly rare, therefore, most mutations perpetuated in a population are neutral. Kimura's neutral theory of molecular evolution has recently been reformulated with the tremendous amount of sequence data that has been derived from investigation into specific genetic systems and genome projects.

Many supposedly neutral mutations have been reevaluated as functional polymorphisms. Polymorphic analysis working with the "relative clinical observed likelihood" (RCOL) has established the relevance of polymorphisms with regards to clinical relevance (Cooper, 2002). Clinically relevant polymorphisms include a change in a functional amino acid, an evolutionarily conserved amino acid, a biochemically non-conserved amino acid and reversion of phenotype with substitution. Although the arginine-lysine substitution does not meet many of the criteria, it does confer with other

characteristics of clinically relevant phenotypes. It segregates with the clinical phenotype, is present in all individuals tested and is not present in two normal dogs from two different breeds and substitutes a conserved amino acid from both mouse and man. The serine-isoleucine substitution also segregates with the affected phenotype but may be more clinically significant. The substitution changes a hydrophilic amino acid for a hydrophobic amino acid. Although not evolutionarily conserved in the human where the corresponding amino acid is an asparagine (another hydrophilic amino acid), it is conserved in the corresponding position in the mouse.

Folding of the colIV molecules into a trimer for incorporation into the basement membrane network is essential for colIV function. ColIV timer assembly begins at the C-terminus and proceeds to the N-terminus. Any disruption in the winding together of these chains can disrupt the normal molecular assembly.

the spectrum of COL4A5 mutations is represented in an investigation of the natural history of 195 XLAS kindreds (Jais et al., 2000). Of the 195 COL4A5 mutations reported, ~20% were large rearrangements, ~15% were microinsertions or microdeletions, ~27% were nonsense mutations and ~15% were splice mutations. Missense mutations were found in 74/195 kindreds and represented ~38% of all mutations in this study. Although the majority of the missense mutations (59/74) were glycine substitutions, 16/74 of the missense mutations were substitutions that involved amino acids other than glycine.

An earlier investigation described many of these non-glycine substitutions in XLAS (Lemmink et al., 1997). Non-glycine substitutions in the COL4A5 gene in this

study were identified as causal mutations in 11/160 kindreds. Two non-glycine substitutions changed a conserved arginine. Six mutations were identified as hydrophobic to hydrophillic changes and three of these changed an isoleucine to a serine. The serine to isoleucine substitution in the COL4A3 gene identified in ECS-HN individuals is a hydrophillic to hydrophobic change, but this report shows that changes in the hydrophobicity of individual amino acids in the COL4A5 gene have been shown to create the XLAS phenotype. Similar changes have yet to be shown in ARAS, but this may be because of the scarcity of reported mutations in these COL4A3 and COL4A4 genes.

While it is more difficult to establish causality of amino acid substitutions on phenotype, we propose that the aberrations identified in the COL4A3 coding region in ECS-HN individuals inhibits the incorporation of the colIVA3 chain into mature colIV trimers. The winding of these molecules could be disrupted by a kink in a collagenous region resulting from a substitution of a hydrophillic amino acid (serine) with a hydrophobic amino acid (isoleucine). The serine-isoleucine subsitution lies adjacent to a glycine molecule and it is possible that the change to a hydrophobic amino acid causes a kink in the chain inhibiting the folding of the glycine into the α3α4α5 helix. Each amino acid interacts with other amino acids in corresponding chains. It is now impossible to determine the intramolecular associations between the α3α4α5 chains at this specific locus, however, the substitution of a larger hydrophobic amino acid at this position may inhibit the folding of these chains into a mature functional trimer.

Another factor inhibiting the winding of the colIVα3α4α5 trimer is the substitution of the conserved arginine near the substituted serine. For whatever reason, this colIVA3 arginine is conserved in mouse and man. These two mutations may act synergistically to inhibit the molecular interactions necessary for trimer assembly.

The ultimate goal of molecular medicine is to be able to predict a clinical phenotype associated with a specific genotype. The identification of a specific base change in ECS COL4A3 and COL4A4 alleles, especially with regards to an amino acid change, may provide such a foundation for the molecular basis of ECS-HN.

REFERENCES

Alport, C.A. 1927. Hereditary Familial Congenital haemorragic nephritis. *The British Medical Journal*. 1:504-506.

Atkin, C.L. 1988. Diseases of the Kidney. Little, Brown, Boston. 617-641 pp.

Atkin, C.L., S.J. Hasstedt, L. Menlove, L. Cannon, N. Kirschner, C. Schwartz, K. Nguyen, and M. Skolnick. 1988. Mapping of Alport syndrome to the long arm of the X chromosome. *American Journal of Human Genetics*. 42:249-55.

Aumailley, M., H. Wiedemann, K. Mann, and R. Timpl. 1989. Binding of nidogen and the laminin-nidogen complex to basement membrane collagen type IV. *European Journal of Biochemistry*. 184:241-8.

Barker, D.F., S.L. Hostikka, J. Zhou, L.T. Chow, A.R. Oliphant, S.C. Gerken, M.C. Gregory, M.H. Skolnick, C.L. Atkin, and K. Tryggvason. 1990. Identification of mutations in the COL4A5 collagen gene in Alport syndrome. *Science*. 248:1224-7.

Battaglia, C., U. Mayer, M. Aumailley, and R. Timpl. 1992. Basement-membrane heparan sulfate proteoglycan binds to laminin by its heparan sulfate chains and to nidogen by sites in the protein core. *European Journal of Biochemistry*. 208:359-66.

Baumal, R., P. Thorner, V.E. Valli, R. McInnes, P. Marrano, R. Jacobs, A. Binnington, and A.G. Bloedow. 1991. Renal disease in carrier female dogs with X-linked hereditary nephritis. Implications for female patients with this disease. *American Journal of Pathology*. 139:751-64.

Beals, R.K., and A.L. Eckhardt. 1969. Hereditary onycho-osteodysplasia (Nail-Patella syndrome). A report of nine kindreds. *Journal of Bone & Joint Surgery - American Volume*. 51:505-16.

Belkin, A.M., and M.A. Stepp. 2000. Integrins as receptors for laminins. *Microscopy Research & Technique*. 51:280-301.

Bellamy, G., and P. Bornstein. 1971. Evidence for procollagen, a biosynthetic precursors of collagen. *Proceedings of the National Academy of Sciences of the United States of America*. 68:1138-42.

Blumberg, B., A.J. MacKrell, and J.H. Fessler. 1988. Drosophila basement membrane procollagen alpha 1(IV). II. Complete cDNA sequence, genomic structure, and general implications for supramolecular assemblies. *Journal of Biological Chemistry*. 263:18328-37.

Blumenthal, S.S., C. Fritsche, and J. Lemann, Jr. 1988. Establishing the diagnosis of benign familial hematuria. The importance of examining the urine sediment of family members. *JAMA*. 259:2263-6.

Boute, N., O. Gribouval, S. Roselli, F. Benessy, H. Lee, A. Fuchshuber, K. Dahan, M.C. Gubler, P. Niaudet, and C. Antignac. 2000. NPHS2, encoding the glomerular protein podocin, is mutated in autosomal recessive steroid-resistant nephrotic syndrome.[erratum appears in *Natature Genetics* 2000 May;25(1):125]. *Nature Genetics*. 24:349-54.

Boye, E., G. Mollet, L. Forestier, L. Cohen-Solal, L. Heidet, P. Cochat, J.P. Grunfeld, J.B. Palcoux, M.C. Gubler, and C. Antignac. 1998. Determination of the genomic structure of the COL4A4 gene and of novel mutations causing autosomal recessive Alport syndrome. *American Journal of Human Genetics*. 63:1329-40.

Boye, E., D. Vetrie, F. Flinter, B. Buckle, T. Pihlajaniemi, E.R. Hamalainen, J.C. Myers, M. Bobrow, and A. Harris. 1991. Major rearrangements in the alpha 5(IV) collagen gene in three patients with Alport syndrome. *Genomics*. 11:1125-32.

Brazel, D., I. Oberbaumer, H. Dieringer, W. Babel, R.W. Glanville, R. Deutzmann, and K. Kuhn. 1987. Completion of the amino acid sequence of the alpha 1 chain of human basement membrane collagen (type IV) reveals 21 non-triplet interruptions located within the collagenous domain. *European Journal of Biochemistry*. 168:529-36.

Brazel, D., R. Pollner, I. Oberbaumer, and K. Kuhn. 1988. Human basement membrane collagen (type IV). The amino acid sequence of the alpha 2(IV) chain and its comparison with the alpha 1(IV) chain reveals deletions in the alpha 1(IV) chain. *European Journal of Biochemistry*. 172:35-42.

Burgeson, R.E., M. Chiquet, R. Deutzmann, P. Ekblom, J. Engel, H. Kleinman, G.R. Martin, G. Meneguzzi, M. Paulsson, and J. Sanes. 1994. A new nomenclature for the laminins. *Matrix Biology*. 14:209-11.

Butkowski, R.J., J.P. Langeveld, J. Wieslander, J. Hamilton, and B.G. Hudson. 1987. Localization of the Goodpasture epitope to a novel chain of basement membrane collagen. *Journal of Biological Chemistry*. 262:7874-7.

Butkowski, R.J., J. Wieslander, M. Kleppel, A.F. Michael, and A.J. Fish. 1989. Basement membrane collagen in the kidney: regional localization of novel chains related to collagen IV. *Kidney International*. 35:1195-202.

Buzza, M., Y.Y. Wang, H. Dagher, J.J. Babon, R.G. Cotton, H. Powell, J. Dowling, and J. Savige. 2001a. COL4A4 mutation in thin basement membrane disease previously described in Alport syndrome.[comment]. *Kidney International*. 60:480-3.

Buzza, M., D. Wilson, and J. Savige. 2001b. Segregation of hematuria in thin basement membrane disease with haplotypes at the loci for Alport syndrome. *Kidney International*. 59:1670-6.

Carlin, B., R. Jaffe, B. Bender, and A.E. Chung. 1981. Entactin, a novel basal lamina-associated sulfated glycoprotein. *Journal of Biological Chemistry*. 256:5209-14.

Caulfield, J.P., and M.G. Farquhar. 1974. The permeability of glomerular capillaries to graded dextrans. Identification of the basement membrane as the primary filtration barrier. *Journal of Cell Biology*. 63:883-903.

Caulfield, J.P., and M.G. Farquhar. 1978. Loss of anionic sites from the glomerular basement membrane in aminonucleoside nephrosis. *Laboratory Investigation*. 39:505-12.

Chance, J.K., and J.A. Stanley. 1977. Alport's syndrome: case report and review of ocular manifestations. *Annals of Ophthalmology*. 9:1527-30.

Chang, R.L., W.M. Deen, C.R. Robertson, and B.M. Brenner. 1975. Permselectivity of the glomerular capillary wall: III. Restricted transport of polyanions. *Kidney International*. 8:212-8.

Chu, M.L., W. de Wet, M. Bernard, J.F. Ding, M. Morabito, J. Myers, C. Williams, and F. Ramirez. 1984. Human pro alpha 1(I) collagen gene structure reveals evolutionary conservation of a pattern of introns and exons. *Nature*. 310:337-40.

Churg, J., and R.L. Sherman. 1973. Pathologic characteristics of hereditary nephritis. *Archives of Pathology & Laboratory Medicine*. 95:374-9.

Cochat, P., P. Guibaud, R. Garcia Torres, B. Roussel, V. Guarner, and F. Larbre. 1988. Diffuse leiomyomatosis in Alport syndrome. *Journal of Pediatrics*. 113:339-43.

Cooper, D.N. 2002. Human gene mutation in pathology and evolution. *Journal of Inherited Metabolic Disease*. 25:157-82.

Cotman, S.L., W. Halfter, and G.J. Cole. 1999. Identification of extracellular matrix ligands for the heparan sulfate proteoglycan agrin. *Experimental Cell Research*. 249:54-64.

Denzer, A.J., R. Brandenberger, M. Gesemann, M. Chiquet, and M.A. Ruegg. 1997. Agrin binds to the nerve-muscle basal lamina via laminin. *Journal of Cell Biology*. 137:671-83.

Denzer, A.J., M. Gesemann, B. Schumacher, and M.A. Ruegg. 1995. An amino-terminal extension is required for the secretion of chick agrin and its binding to extracellular matrix. *Journal of Cell Biology*. 131:1547-60.

Denzer, A.J., T. Schulthess, C. Fauser, B. Schumacher, R.A. Kammerer, J. Engel, and M.A. Ruegg. 1998. Electron microscopic structure of agrin and mapping of its binding site in laminin-1. *EMBO Journal*. 17:335-43.

DeSimone, D.W., M.A. Stepp, R.S. Patel, and R.O. Hynes. 1987. The integrin family of cell surface receptors. *Biochemical Society Transactions*. 15:789-91.

Ding, J., J. Stitzel, P. Berry, E. Hawkins, and C.E. Kashtan. 1995. Autosomal recessive Alport syndrome: mutation in the COL4A3 gene in a woman with Alport syndrome and posttransplant antiglomerular basement membrane nephritis. *Journal of the American Society of Nephrology*. 5:1714-7.

Dixit, S.N., J.M. Stuart, J.M. Seyer, J. Risteli, R. Timpl, and A.H. Kang. 1981. Type IV collagens' isolation and characterization of 7S collagen from human kidney, liver and lung. *Collagen & Related Research*. 1:549-56.

Dolz, R., J. Engel, and K. Kuhn. 1988. Folding of collagen IV. *European Journal of Biochemistry*. 178:357-66.

Dreyer, S.D., G. Zhou, A. Baldini, A. Winterpacht, B. Zabel, W. Cole, R.L. Johnson, and B. Lee. 1998. Mutations in LMX1B cause abnormal skeletal patterning and renal dysplasia in nail patella syndrome. *Nature Genetics*. 19:47-50.

Durkin, M.E., S. Chakravarti, B.B. Bartos, S.H. Liu, R.L. Friedman, and A.E. Chung. 1988. Amino acid sequence and domain structure of entactin. Homology with epidermal growth factor precursor and low density lipoprotein receptor. *Journal of Cell Biology*. 107:2749-56.

Dziadek, M., and R. Timpl. 1985. Expression of nidogen and laminin in basement membranes during mouse embryogenesis and in teratocarcinoma cells. *Developmental Biology*. 111:372-82.

Epstein, C.J., M.A. Sahud, C.F. Piel, J.R. Goodman, M.R. Bernfield, J.H. Kushner, and A.R. Ablin. 1972. Hereditary macrothrombocytopathia, nephritis and deafness. *American Journal of Medicine*. 52:299-310.

Erickson, A.C., and J.R. Couchman. 2000. Still more complexity in mammalian basement membranes. *Journal of Histochemistry & Cytochemistry*. 48:1291-306.

Farquhar, M.G., S.L. Wissig, and G.E. Palade. 1961. Glomerular pemability: I. Ferritin transfer across the normal glomerular capillary wall. *Journal of Experimental Medicine*. 113:47-66.

Feingold, J., E. Bois, A. Chompret, M. Broyer, M.C. Gubler, and J.P. Grunfeld. 1985. Genetic heterogeneity of Alport syndrome. *Kidney International*. 27:672-7.

Fox, J.W., U. Mayer, R. Nischt, M. Aumailley, D. Reinhardt, H. Wiedemann, K. Mann, R. Timpl, T. Krieg, and J. Engel. 1991. Recombinant nidogen consists of three globular domains and mediates binding of laminin to collagen type IV. *EMBO Journal*. 10:3137-46.

Fuchshuber, A., G. Jean, O. Gribouval, M.C. Gubler, M. Broyer, J.S. Beckmann, P. Niaudet, and C. Antignac. 1995. Mapping a gene (SRN1) to chromosome 1q25-q31 in idiopathic nephrotic syndrome confirms a distinct entity of autosomal recessive nephrosis. *Human Molecular Genetics*. 4:2155-8.

Gautam, M., P.G. Noakes, L. Moscoso, F. Rupp, R.H. Scheller, J.P. Merlie, and J.R. Sanes. 1996. Defective neuromuscular synaptogenesis in agrin-deficient mutant mice. *Cell*. 85:525-35.

Gauthier, B., H. Trachtman, R. Frank, and E. Valderrama. 1989. Familial thin basement membrane nephropathy in children with asymptomatic microhematuria. *Nephron*. 51:502-8.

Govan, J.A. 1983. Ocular manifestations of Alport's syndrome: a hereditary disorder of basement membranes? *British Journal of Ophthalmology*. 67:493-503.

Graham, R.C., Jr., and M.J. Karnovsky. 1966. Glomerular permeability. Ultrastructural cytochemical studies using peroxidases as protein tracers. *Journal of Experimental Medicine*. 124:1123-34.

Grant, D.S., C.P. Leblond, H.K. Kleinman, S. Inoue, and J.R. Hassell. 1989. The incubation of laminin, collagen IV, and heparan sulfate proteoglycan at 35 degrees C yields basement membrane-like structures. *Journal of Cell Biology*. 108:1567-74.

Griffin, C.A., B.S. Emanuel, J.R. Hansen, W.K. Cavenee, and J.C. Myers. 1987. Human collagen genes encoding basement membrane alpha 1 (IV) and alpha 2 (IV) chains map to the distal long arm of chromosome 13. *Proceedings of the National Academy of Sciences of the United States of America*. 84:512-6.

Groffen, A.J., F.W. Hop, K. Tryggvason, H. Dijkman, K.J. Assmann, J.H. Veerkamp, L.A. Monnens, and L.P. Van den Heuvel. 1997. Evidence for the existence of multiple heparan sulfate proteoglycans in the human glomerular basement membrane and mesangial matrix. *European Journal of Biochemistry*. 247:175-82.

Groffen, A.J., M.A. Ruegg, H. Dijkman, T.J. van de Velden, C.A. Buskens, J. van den Born, K.J. Assmann, L.A. Monnens, J.H. Veerkamp, and L.P. van den Heuvel. 1998. Agrin is a major heparan sulfate proteoglycan in the human glomerular basement membrane. *Journal of Histochemistry & Cytochemistry*. 46:19-27.

Grunfeld, J.P. 1985. The clinical spectrum of hereditary nephritis. *Kidney International*. 27:83-92.

Gubler, M., M. Levy, M. Broyer, C. Naizot, G. Gonzales, D. Perrin, and R. Habib. 1981. Alport's syndrome. A report of 58 cases and a review of the literature. *American Journal of Medicine*. 70:493-505.

Gubler, M.C., B. Knebelmann, A. Beziau, M. Broyer, Y. Pirson, F. Haddoum, M.M. Kleppel, and C. Antignac. 1995. Autosomal recessive Alport syndrome: immunohistochemical study of type IV collagen chain distribution. *Kidney International*. 47:1142-7.

Gubler, M.C., M. Levy, C. Naizot, and R. Habib. 1980. Glomerular basement membrane changes in hereditary glomerular diseases. *Renal Physiology*. 3:405-13.

Gunwar, S., F. Ballester, M.E. Noelken, Y. Sado, Y. Ninomiya, and B.G. Hudson. 1998. Glomerular basement membrane. Identification of a novel disulfide-cross-linked network of alpha3, alpha4, and alpha5 chains of type IV collagen and its implications for the pathogenesis of Alport syndrome. *Journal of Biological Chemistry*. 273:8767-75.

Guyton, A.C. 1991. Textbook of Medical Physiology. W.B. Saunders, New York.

Halfter, W., S. Dong, B. Schurer, and G.J. Cole. 1998. Collagen XVIII is a basement membrane heparan sulfate proteoglycan. *Journal of Biological Chemistry*. 273:25404-12.

Haralson, M.A., S.J. Federspiel, A. Martinez-Hernandez, R.K. Rhodes, and E.J. Miller. 1985. Synthesis of [pro alpha 1(IV)]3 collagen molecules by cultured embryo-derived parietal yolk sac cells. *Biochemistry*. 24:5792-7.

Heidet, L., C. Arrondel, L. Forestier, L. Cohen-Solal, G. Mollet, B. Gutierrez, C. Stavrou, M.C. Gubler, and C. Antignac. 2001. Structure of the human type IV collagen gene COL4A3 and mutations in autosomal Alport syndrome. *Journal of the American Society of Nephrology*. 12:97-106.

Heptinstall, R.H. 1992. Pathology of the Kidney. Little, Brown, Boston.

Hinglais, N., J.P. Grunfeld, and E. Bois. 1972. Characteristic ultrastructural lesion of the glomerular basement membrane in progressive hereditary nephritis (Alport's syndrome). *Laboratory Investigation*. 27:473-87.

Hofmann, H., T. Voss, K. Kuhn, and J. Engel. 1984. Localization of flexible sites in thread-like molecules from electron micrographs. Comparison of interstitial, basement membrane and intima collagens. *Journal of Molecular Biology*. 172:325-43.

Holzman, L.B., P.L. St John, I.A. Kovari, R. Verma, H. Holthofer, and D.R. Abrahamson. 1999. Nephrin localizes to the slit pore of the glomerular epithelial cell. *Kidney International*. 56:1481-91.

Hood, J.C., C. Huxtable, I. Naito, C. Smith, R. Sinclair, and J. Savige. 2002. A novel model of autosomal dominant Alport syndrome in Dalmatian dogs. *Nephrology Dialysis Transplantation*. 17:2094-8.

Hood, J.C., W.F. Robinson, C.R. Huxtable, J.S. Bradley, R.J. Sutherland, and M.A. Thomas. 1990. Hereditary nephritis in the bull terrier: evidence for inheritance by an autosomal dominant gene. *Veterinary Record*. 126:456-9.

Hood, J.C., J. Savige, A. Hendtlass, M.M. Kleppel, C.R. Huxtable, and W.F. Robinson. 1995. Bull terrier hereditary nephritis: a model for autosomal dominant Alport syndrome. *Kidney International*. 47:758-65.

Hostikka, S.L., R.L. Eddy, M.G. Byers, M. Hoyhtya, T.B. Shows, and K. Tryggvason. 1990. Identification of a distinct type IV collagen alpha chain with restricted kidney distribution and assignment of its gene to the locus of X chromosome-linked Alport syndrome. *Proceedings of the National Academy of Sciences of the United States of America*. 87:1606-10.

Hostikka, S.L., and K. Tryggvason. 1988. The complete primary structure of the alpha 2 chain of human type IV collagen and comparison with the alpha 1(IV) chain. *Journal of Biological Chemistry*. 263:19488-93.

Hudson, B.G., S.T. Reeders, and K. Tryggvason. 1993. Type IV collagen: structure, gene organization, and role in human diseases. Molecular basis of Goodpasture and Alport syndromes and diffuse leiomyomatosis. *Journal of Biological Chemistry*. 268:26033-6.

Humphries, M.J. 2000. Integrin structure. *Biochemical Society Transactions*. 28:311-39.

Ichikawa, I., and A. Fogo. 1996. Focal segmental glomerulosclerosis.[comment]. *Pediatric Nephrology*. 10:374-91.

Jais, J.P., B. Knebelmann, I. Giatras, M. De Marchi, G. Rizzoni, A. Renieri, M. Weber, O. Gross, K.O. Netzer, F. Flinter, Y. Pirson, C. Verellen, J. Wieslander, U. Persson, K. Tryggvason, P. Martin, J.M. Hertz, C. Schroder, M. Sanak, S. Krejcova, M.F. Carvalho, J. Saus, C. Antignac, H. Smeets, and M.C. Gubler. 2000. X-linked Alport syndrome: natural history in 195 families and genotype-phenotype correlations in males. *Journal of the American Society of Nephrology*. 11:649-57.

Jansen, B., P. Thorner, R. Baumal, V. Valli, M.G. Maxie, and A. Singh. 1986a. Samoyed hereditary glomerulopathy (SHG). Evolution of splitting of glomerular capillary basement membranes. *American Journal of Pathology*. 125:536-45.

Jansen, B., P.S. Thorner, A. Singh, J.M. Patterson, J.H. Lumsden, V.E. Valli, R. Baumal, and R.K. Basrur. 1984. Animal model of human disease: hereditary nephritis in Samoyed dogs. *American Journal of Pathology*. 116:175-8.

Jansen, B., L. Tryphonas, J. Wong, P. Thorner, M.G. Maxie, V.E. Valli, R. Baumal, and P.K. Basrur. 1986b. Mode of inheritance of Samoyed hereditary glomerulopathy: an animal model for hereditary nephritis in humans. *Journal of Laboratory & Clinical Medicine*. 107:551-5.

Jefferson, J.A., H.H. Lemmink, A.E. Hughes, C.M. Hill, H.J. Smeets, C.C. Doherty, and A.P. Maxwell. 1997. Autosomal dominant Alport syndrome linked to the type IV collage alpha 3 and alpha 4 genes (COL4A3 and COL4A4). *Nephrology Dialysis Transplantation*. 12:1595-9.

Johansson, C., R. Butkowski, and J. Wieslander. 1992. The structural organization of type IV collagen. Identification of three NC1 populations in the glomerular basement membrane. *Journal of Biological Chemistry*. 267:24533-7.

Kahsai, T.Z., G.C. Enders, S. Gunwar, C. Brunmark, J. Wieslander, R. Kalluri, J. Zhou, M.E. Noelken, and B.G. Hudson. 1997. Seminiferous tubule basement membrane. Composition and organization of type IV collagen chains, and the linkage of alpha3(IV) and alpha5(IV) chains. *Journal of Biological Chemistry*. 272:17023-32.

Kallunki, P., R.L. Eddy, M.G. Byers, M. Kestila, T.B. Shows, and K. Tryggvason. 1991. Cloning of human heparan sulfate proteoglycan core protein, assignment of the gene (HSPG2) to 1p36.1----p35 and identification of a BamHI restriction fragment length polymorphism. *Genomics*. 11:389-96.

Kalluri, R., C.F. Shield, P. Todd, B.G. Hudson, and E.G. Neilson. 1997. Isoform switching of type IV collagen is developmentally arrested in X-linked Alport syndrome leading to increased susceptibility of renal basement membranes to endoproteolysis. *Journal of Clinical Investigation*. 99:2470-8.

Kamagata, Y., M.G. Mattei, and Y. Ninomiya. 1992. Isolation and sequencing of cDNAs and genomic DNAs encoding the alpha 4 chain of basement membrane collagen type IV and assignment of the gene to the distal long arm of human chromosome 2. *Journal of Biological Chemistry*. 267:23753-8.

Kanwar, Y.S., and M.G. Farquhar. 1979a. Anionic sites in the glomerular basement membrane. In vivo and in vitro localization to the laminae rarae by cationic probes. *Journal of Cell Biology*. 81:137-53.

Kanwar, Y.S., and M.G. Farquhar. 1979b. Presence of heparan sulfate in the glomerular basement membrane. *Proceedings of the National Academy of Sciences of the United States of America*. 76:1303-7.

Kao, W.W., D.J. Prockop, and R.A. Berg. 1979. Kinetics for the secretion of nonhelical procollagen by freshly isolated tendon cells. *Journal of Biological Chemistry*. 254:2234-43.

Kaplan, J.M., S.H. Kim, K.N. North, H. Rennke, L.A. Correia, H.Q. Tong, B.J. Mathis, J.C. Rodriguez-Perez, P.G. Allen, A.H. Beggs, and M.R. Pollak. 2000. Mutations in ACTN4, encoding alpha-actinin-4, cause familial focal segmental glomerulosclerosis. *Nature Genetics*. 24:251-6.

Kashtan, C.E., M.M. Kleppel, and M.C. Gubler. 1996. Immunohistologic findings in Alport syndrome. *Contributions to Nephrology*. 117:142-53.

Kawai, S., S. Nomura, T. Harano, K. Harano, T. Fukushima, and G. Osawa. 1996. The COL4A5 gene in Japanese Alport syndrome patients: spectrum of mutations of all exons. The Japanese Alport Network. *Kidney International*. 49:814-22.

Kefalides, N.A. 1971. Isolation of a collagen from basement membranes containing three identical - chains. *Biochemical & Biophysical Research Communications*. 45:226-34.

Kestila, M., U. Lenkkeri, M. Mannikko, J. Lamerdin, P. McCready, H. Putaala, V. Ruotsalainen, T. Morita, M. Nissinen, R. Herva, C.E. Kashtan, L. Peltonen, C. Holmberg, A. Olsen, and K. Tryggvason. 1998. Positionally cloned gene for a novel glomerular protein--nephrin--is mutated in congenital nephrotic syndrome. *Molecular Cell*. 1:575-82.

Kimura, N., T. Toyoshima, T. Kojima, and M. Shimane. 1998. Entactin-2: a new member of basement membrane protein with high homology to entactin/nidogen. *Experimental Cell Research*. 241:36-45.

Kivirikko, K.I., and R. Myllyla. 1987. Recent developments in posttranslational modification: intracellular processing. *Methods in Enzymology*. 144:96-114.

Kleinman, H.K., M.L. McGarvey, J.R. Hassell, V.L. Star, F.B. Cannon, G.W. Laurie, and G.R. Martin. 1986. Basement membrane complexes with biological activity. *Biochemistry*. 25:312-8.

Kleppel, M.M., W.W. Fan, H.I. Cheong, C.E. Kashtan, and A.F. Michael. 1992. Immunochemical studies of the Alport antigen. *Kidney International*. 41:1629-37.

Kleppel, M.M., C. Kashtan, P.A. Santi, J. Wieslander, and A.F. Michael. 1989a. Distribution of familial nephritis antigen in normal tissue and renal basement membranes of patients with homozygous and heterozygous Alport familial nephritis. Relationship of familial nephritis and Goodpasture antigens to novel collagen chains and type IV collagen. *Laboratory Investigation*. 61:278-89.

Kleppel, M.M., P.A. Santi, J.D. Cameron, J. Wieslander, and A.F. Michael. 1989b. Human tissue distribution of novel basement membrane collagen. *American Journal of Pathology*. 134:813-25.

Kohfeldt, E., T. Sasaki, W. Gohring, and R. Timpl. 1998. Nidogen-2: a new basement membrane protein with diverse binding properties. *Journal of Molecular Biology*. 282:99-109.

Kuhn, K., H. Wiedemann, R. Timpl, J. Risteli, H. Dieringer, T. Voss, and R.W. Glanville. 1981. Macromolecular structure of basement membrane collagens. *FEBS Letters*. 125:123-8.

Layman, D.L., E.B. McGoodwin, and G.R. Martin. 1971. The nature of the collagen synthesized by cultured human fibroblasts. *Proceedings of the National Academy of Sciences of the United States of America*. 68:454-8.

Leblond, C.P., and S. Inoue. 1989. Structure, composition, and assembly of basement membrane. *American Journal of Anatomy*. 185:367-90.

Lees, G.E., R.G. Helman, L.D. Homco, N.J. Millichamp, J.F. Hunter, and M.S. Frey. 1998a. Early diagnosis of familial nephropathy in english cocker spaniels. *Journal of the American Animal Hospital Association*. 34:189-195.

Lees, G.E., R.G. Helman, C.E. Kashtan, A.F. Michael, L.D. Homco, N.J. Millichamp, Z.T. Camacho, J.W. Templeton, Y. Ninomiya, Y. Sado, I. Naito, and Y. Kim. 1999. New form of X-linked dominant hereditary nephritis in dogs. *American Journal of Veterinary Research*. 60:373-83.

Lees, G.E., R.G. Helman, C.E. Kashtan, A.F. Michael, L.D. Homco, N.J. Millichamp, Y. Ninomiya, Y. Sado, I. Naito, and Y. Kim. 1998b. A Model of autosomal recessive alport-syndrome in english cocker spaniel dogs. *Kidney International*. 54:706-719.

Lees, G.E., C.E. Kashtan, A.F. Michael, R.G. Helman, I. Naito, Y. Ninomiya, Y. Sado, and Y. Kim. 2000. Expression of the alpha6 chain of type IV collagen in glomerular basement membranes of healthy adult dogs. *American Journal of Veterinary Research*. 61:38-41.

Lees, G.E., P.D. Wilson, R.G. Helman, L.D. Homco, and M.S. Frey. 1997. Glomerular ultrastructural findings similar to hereditary nephritis in 4 English cocker spaniels. *Journal of Veterinary Internal Medicine*. 11:80-5.

Leinonen, A., M. Mariyama, T. Mochizuki, K. Tryggvason, and S.T. Reeders. 1994. Complete primary structure of the human type IV collagen alpha 4(IV) chain. Comparison with structure and expression of the other alpha (IV) chains. *Journal of Biological Chemistry*. 269:26172-7.

Lemmink, H.H., T. Mochizuki, L.P. van den Heuvel, C.H. Schroder, A. Barrientos, L.A. Monnens, B.A. van Oost, H.G. Brunner, S.T. Reeders, and H.J. Smeets. 1994. Mutations in the type IV collagen alpha 3 (COL4A3) gene in autosomal recessive Alport syndrome. *Human Molecular Genetics*. 3:1269-73.

Lemmink, H.H., W.N. Nillesen, T. Mochizuki, C.H. Schroder, H.G. Brunner, B.A. van Oost, L.A. Monnens, and H.J. Smeets. 1996. Benign familial hematuria due to mutation of the type IV collagen alpha4 gene. *Journal of Clinical Investigation*. 98:1114-8.

Lemmink, H.H., C.H. Schroder, L.A. Monnens, and H.J. Smeets. 1997. The clinical spectrum of type IV collagen mutations. *Human Mutation*. 9:477-99.

Lenkkeri, U., M. Mannikko, P. McCready, J. Lamerdin, O. Gribouval, P.M. Niaudet, C.K. Antignac, C.E. Kashtan, C. Homberg, A. Olsen, M. Kestila, and K. Tryggvason. 1999. Structure of the gene for congenital nephrotic syndrome of the finnish type (NPHS1) and characterization of mutations. *American Journal of Human Genetics*. 64:51-61.

Longo, I., P. Porcedda, F. Mari, D. Giachino, I. Meloni, C. Deplano, A. Brusco, M. Bosio, L. Massella, G. Lavoratti, D. Roccatello, G. Frasca, G. Mazzucco, A.O. Muda, M. Conti, F. Fasciolo, C. Arrondel, L. Heidet, A. Renieri, and M. De Marchi. 2002. COL4A3/COL4A4 mutations: from familial hematuria to autosomal-dominant or recessive Alport syndrome. *Kidney International*. 61:1947-56.

Macdougall, D.F., T. Cook, A.P. Steward, and V. Cattell. 1986. Canine chronic renal disease: prevalence and types of glomerulonephritis in the dog. *Kidney International*. 29:1144-51.

Mann, K., R. Deutzmann, M. Aumailley, R. Timpl, L. Raimondi, Y. Yamada, T.C. Pan, D. Conway, and M.L. Chu. 1989. Amino acid sequence of mouse nidogen, a multidomain basement membrane protein with binding activity for laminin, collagen IV and cells. *EMBO Journal*. 8:65-72.

Mann, K., R. Deutzmann, and R. Timpl. 1988. Characterization of proteolytic fragments of the laminin-nidogen complex and their activity in ligand-binding assays. *European Journal of Biochemistry*. 178:71-80.

Mariyama, M., A. Leinonen, T. Mochizuki, K. Tryggvason, and S.T. Reeders. 1994. Complete primary structure of the human alpha 3(IV) collagen chain. Coexpression of the alpha 3(IV) and alpha 4(IV) collagen chains in human tissues. *Journal of Biological Chemistry*. 269:23013-7.

Mariyama, M., K. Zheng, T.L. Yang-Feng, and S.T. Reeders. 1992. Colocalization of the genes for the alpha 3(IV) and alpha 4(IV) chains of type IV collagen to chromosome 2 bands q35-q37. *Genomics*. 13:809-13.

Martin, P., N. Heiskari, H. Pajari, C. Gronhagen-Riska, H. Kaariainen, O. Koskimies, and K. Tryggvason. 2000. Spectrum of COL4A5 mutations in Finnish Alport syndrome patients. *Human Mutation*. 15:579.

Martin, P., N. Heiskari, J. Zhou, A. Leinonen, T. Tumelius, J.M. Hertz, D. Barker, M. Gregory, C. Atkin, U. Styrkarsdottir, H. Neumann, J. Springate, T. Shows, E. Pettersson, and K. Tryggvason. 1998. High mutation detection rate in the COL4A5 collagen gene in suspected Alport syndrome using PCR and direct DNA sequencing. *Journal of the American Society of Nephrology*. 9:2291-301.

Martin, P.T., and J.R. Sanes. 1997. Integrins mediate adhesion to agrin and modulate agrin signaling. *Development*. 124:3909-17.

Mayer, U., R. Nischt, E. Poschl, K. Mann, K. Fukuda, M. Gerl, Y. Yamada, and R. Timpl. 1993. A single EGF-like motif of laminin is responsible for high affinity nidogen binding. *EMBO Journal*. 12:1879-85.

Miller, E.J., E.H. Epstein, Jr., and K.A. Piez. 1971. Identification of three genetically distinct collagens by cyanogen bromide cleavage of insoluble human skin and cartilage collagen. *Biochemical & Biophysical Research Communications*. 42:1024-9.

Miller, E.J., and V.J. Matukas. 1969. Chick cartilage collagen: a new type of alpha 1 chain not present in bone or skin of the species. *Proceedings of the National Academy of Sciences of the United States of America*. 64:1264-8.

Miner, J.H., J. Cunningham, and J.R. Sanes. 1998. Roles for laminin in embryogenesis: exencephaly, syndactyly, and placentopathy in mice lacking the laminin alpha5 chain. *Journal of Cell Biology*. 143:1713-23.

Miner, J.H., and C. Li. 2000. Defective glomerulogenesis in the absence of laminin alpha5 demonstrates a developmental role for the kidney glomerular basement membrane. *Developmental Biology*. 217:278-89.

Miner, J.H., R. Morello, K.L. Andrews, C. Li, C. Antignac, A.S. Shaw, and B. Lee. 2002. Transcriptional induction of slit diaphragm genes by Lmx1b is required in podocyte differentiation. *Journal of Clinical Investigation*. 109:1065-72.

Miner, J.H., B.L. Patton, S.I. Lentz, D.J. Gilbert, W.D. Snider, N.A. Jenkins, N.G. Copeland, and J.R. Sanes. 1997. The laminin alpha chains: expression, developmental transitions, and chromosomal locations of alpha1-5, identification of heterotrimeric laminins 8-11, and cloning of a novel alpha3 isoform. *Journal of Cell Biology*. 137:685-701.

Mochizuki, T., H.H. Lemmink, M. Mariyama, C. Antignac, M.C. Gubler, Y. Pirson, C. Verellen-Dumoulin, B. Chan, C.H. Schroder, and H.J. Smeets. 1994. Identification of mutations in the alpha 3(IV) and alpha 4(IV) collagen genes in autosomal recessive Alport syndrome. *Nature Genetics*. 8:77-81.

Moeller, M.J., I.A. Kovari, and L.B. Holzman. 2000. Evaluation of a new tool for exploring podocyte biology: mouse Nphs1 5' flanking region drives LacZ expression in podocytes. *Journal of the American Society of Nephrology*. 11:2306-14.

Morello, R., G. Zhou, S.D. Dreyer, S.J. Harvey, Y. Ninomiya, P.S. Thorner, J.H. Miner, W. Cole, A. Winterpacht, B. Zabel, K.C. Oberg, and B. Lee. 2001. Regulation of glomerular basement membrane collagen expression by LMX1B contributes to renal disease in nail patella syndrome. *Nature Genetics*. 27:205-8.

Morrison, K.E., M. Mariyama, T.L. Yang-Feng, and S.T. Reeders. 1991. Sequence and localization of a partial cDNA encoding the human alpha 3 chain of type IV collagen. *American Journal of Human Genetics*. 49:545-54.

Muragaki, Y., S. Timmons, C.M. Griffith, S.P. Oh, B. Fadel, T. Quertermous, and B.R. Olsen. 1995. Mouse Col18a1 is expressed in a tissue-specific manner as three alternative variants and is localized in basement membrane zones. *Proceedings of the National Academy of Sciences of the United States of America*. 92:8763-7.

Murray, M., and N.C. Wright. 1974. A morphologic study of canine glomerulonephritis. *Laboratory Investigation.* 30:213-21.

Nagayoshi, T., D. Sanborn, N.J. Hickok, D.R. Olsen, M.J. Fazio, M.L. Chu, R. Knowlton, K. Mann, R. Deutzmann, and R. Timpl. 1989. Human nidogen: complete amino acid sequence and structural domains deduced from cDNAs, and evidence for polymorphism of the gene. *DNA*. 8:581-94.

Nakanishi, K., N. Yoshikawa, K. Iijima, K. Kitagawa, H. Nakamura, H. Ito, K. Yoshioka, M. Kagawa, and Y. Sado. 1994. Immunohistochemical study of alpha 1-5 chains of type IV collagen in hereditary nephritis. *Kidney International*. 46:1413-21.

Netzer, K.O., A. Leinonen, A. Boutaud, D.B. Borza, P. Todd, S. Gunwar, J.P. Langeveld, and B.G. Hudson. 1999a. The goodpasture autoantigen. Mapping the major conformational epitope(s) of alpha3(IV) collagen to residues 17-31 and 127-141 of the NC1 domain. *Journal of Biological Chemistry*. 274:11267-74.

Netzer, K.O., S. Seibold, and M. Weber. 1999b. Thin basement membrane--do we have a window for understanding the molecular pathogenesis? *Nephrology Dialysis Transplantation*. 14:1060-1.

Nielsen, C.E. 1978. Lenticonus anterior and Alport's syndrome. *Acta Ophthalmologica*. 56:518-30.

Nitkin, R.M., M.A. Smith, C. Magill, J.R. Fallon, Y.M. Yao, B.G. Wallace, and U.J. McMahan. 1987. Identification of agrin, a synaptic organizing protein from Torpedo electric organ. *Journal of Cell Biology*. 105:2471-8.

Noakes, P.G., J.H. Miner, M. Gautam, J.M. Cunningham, J.R. Sanes, and J.P. Merlie. 1995. The renal glomerulus of mice lacking s-laminin/laminin beta 2: nephrosis despite molecular compensation by laminin beta 1. *Nature Genetics*. 10:400-6.

Oh, S.P., Y. Kamagata, Y. Muragaki, S. Timmons, A. Ooshima, and B.R. Olsen. 1994. Isolation and sequencing of cDNAs for proteins with multiple domains of Gly-Xaa-Yaa repeats identify a distinct family of collagenous proteins. *Proceedings of the National Academy of Sciences of the United States of America*. 91:4229-33.

Olsen, D.R., T. Nagayoshi, M. Fazio, M.G. Mattei, E. Passage, D. Weil, R. Timpl, M.L. Chu, and J. Uitto. 1989. Human nidogen: cDNA cloning, cellular expression, and mapping of the gene to chromosome Iq43. *American Journal of Human Genetics*. 44:876-85.

Oohashi, T., M. Sugimoto, M.G. Mattei, and Y. Ninomiya. 1994. Identification of a new collagen IV chain, alpha 6(IV), by cDNA isolation and assignment of the gene to chromosome Xq22, which is the same locus for COL4A5. *Journal of Biological Chemistry*. 269:7520-6.

Ozen, S., D. Ertoy, L. Heidet, L. Cohen-Solal, H. Ozen, N. Besbas, A. Bakkaoglu, and C. Antignac. 2001. Benign familial hematuria associated with a novel COL4A4 mutation. *Pediatric Nephrology*. 16:874-7.

Paulsson, M., P.D. Yurchenco, G.C. Ruben, J. Engel, and R. Timpl. 1987. Structure of low density heparan sulfate proteoglycan isolated from a mouse tumor basement membrane. *Journal of Molecular Biology*. 197:297-313.

Pease, D.C., and R.F. Baker. 1950. Electron microscopy of the kidney. *American Journal of Anatomy*. 87:349-389.

Peissel, B., L. Geng, R. Kalluri, C. Kashtan, H.G. Rennke, G.R. Gallo, K. Yoshioka, M.J. Sun, B.G. Hudson, and E.G. Neilson. 1995. Comparative distribution of the alpha 1(IV), alpha 5(IV), and alpha 6(IV) collagen chains in normal human adult and fetal tissues and in kidneys from X-linked Alport syndrome patients. *Journal of Clinical Investigation*. 96:1948-57.

Piez, K.A., Eigner, E. A., and Lewis, M.S. 1963. The chromatographic separation and amino acid composition of several collagens. *biochemistry*. 2:58-66.

Pihlajaniemi, T., E.R. Pohjolainen, and J.C. Myers. 1990. Complete primary structure of the triple-helical region and the carboxyl-terminal domain of a new type IV collagen chain, alpha 5(IV). *Journal of Biological Chemistry*. 265:13758-66.

Pinnell, S.R., S.M. Krane, J.E. Kenzora, and M.J. Glimcher. 1972. A heritable disorder of connective tissue. Hydroxylysine-deficient collagen disease. *New England Journal of Medicine*. 286:1013-20.

Pinnell, S.R., and G.R. Martin. 1968. The cross-linking of collagen and elastin: enzymatic conversion of lysine in peptide linkage to alpha-aminoadipic-delta-semialdehyde (allysine) by an extract from bone. *Proceedings of the National Academy of Sciences of the United States of America*. 61:708-16.

Pochet, J.M., G. Bobrie, P. Landais, B. Goldfarb, and J.P. Grunfeld. 1989. Renal prognosis in Alport's and related syndromes: influence of the mode of inheritance.[comment]. *Nephrology Dialysis Transplantation*. 4:1016-21.

Poschl, E., R. Pollner, and K. Kuhn. 1988. The genes for the alpha 1(IV) and alpha 2(IV) chains of human basement membrane collagen type IV are arranged head-to-head and separated by a bidirectional promoter of unique structure. *EMBO Journal*. 7:2687-95.

Prockop, D.J., and K.I. Kivirikko. 1995. Collagens: molecular biology, diseases, and potentials for therapy. *Annual Review of Biochemistry*. 64:403-34.

Putaala, H., K. Sainio, H. Sariola, and K. Tryggvason. 2000. Primary structure of mouse and rat nephrin cDNA and structure and expression of the mouse gene. *Journal of the American Society of Nephrology*. 11:991-1001.

Putaala, H., R. Soininen, P. Kilpelainen, J. Wartiovaara, and K. Tryggvason. 2001. The murine nephrin gene is specifically expressed in kidney, brain and pancreas: inactivation of the gene leads to massive proteinuria and neonatal death. *Human Molecular Genetics*. 10:1-8.

Raats, C.J., M.A. Bakker, W. Hoch, W.P. Tamboer, A.J. Groffen, L.P. van den Heuvel, J.H. Berden, and J. van den Born. 1998. Differential expression of agrin in renal basement membranes as revealed by domain-specific antibodies. *Journal of Biological Chemistry*. 273:17832-8.

Ramachandran, G.N. 1954. Structure of collagen. *Nature*. 174:269-270.

Ramachandran, G.N. 1955. Structure of collagen. *Nature*. 176:593-595.

Rambourg, A., and C.P. Leblond. 1967. Staining of basement membranes and assocated structures by the periodic acid-Schiff and periodicacid-silver methenamine techniques. *Journal of Ultrastructure Research*. 20:306-9.

Reeders, S.T. 1992. Molecular genetics of hereditary nephritis. *Kidney International*. 42:783-92.

Renieri, A., M. Seri, J.C. Myers, T. Pihlajaniemi, A. Sessa, G. Rizzoni, and M. De Marchi. 1992. Alport syndrome caused by a 5' deletion within the COL4A5 gene. *Human Genetics*. 89:120-1.

Rhys, C., B. Snyers, and Y. Pirson. 1997. Recurrent corneal erosion associated with Alport's syndrome. Rapid communication. *Kidney International*. 52:208-11.

Rich, A. 1955. The structure of collagen. *Nature*. 176:915-916.

Risteli, J., G. Wick, and R. Timpl. 1981. Immunological characterization of the 7-S domain of type IV collagens. *Collagen & Related Research*. 1:419-32.

Robinson, W.F., C.R. Huxtable, and J.P. Gooding. 1985. Familial nephropathy in cocker spaniels. *Australian Veterinary Journal*. 62:109-12.

Robinson, W.F., S.E. Shaw, B. Stanley, C.R. Huxtable, A.D. Watson, S.E. Friend, and R. Mitten. 1989. Chronic renal disease in bull terriers. *Australian Veterinary Journal*. 66:193-5.

Rodewald, R., and M.J. Karnovsky. 1974. Porous substructure of the glomerular slit diaphragm in the rat and mouse. *Journal of Cell Biology*. 60:423-33.

Rogers, P.W., N.A. Kurtzman, S.M. Bunn, Jr., and M.G. White. 1973. Familial benign essential hematuria. *Archives of Internal Medicine*. 131:257-62.

Rohrbach, D.H., C.W. Wagner, V.L. Star, G.R. Martin, K.S. Brown, and J.W. Yoon. 1983. Reduced synthesis of basement membrane heparan sulfate proteoglycan in streptozotocin-induced diabetic mice. *Journal of Biological Chemistry*. 258:11672-7.

Rohrbach, D.H.a.T., R. 1991. Molecular and Celular Aspects of Basement Membranes. Academic Press, New York

Rosenzweig, L.J., and Y.S. Kanwar. 1982. Removal of sulfated (heparan sulfate) or nonsulfated (hyaluronic acid) glycosaminoglycans results in increased permeability of the glomerular basement membrane to 125I-bovine serum albumin. *Laboratory Investigation*. 47:177-84.

Rumpelt, H.J., K.H. Langer, K. Scharer, E. Straub, and W. Thoenes. 1974. Split and extremely thin glomerular basement membranes in hereditary nephropathy (Alport's syndrome). *Virchows Archiv. A, Pathological Anatomy & Histology*. 364:225-33.

Ruotsalainen, V., P. Ljungberg, J. Wartiovaara, U. Lenkkeri, M. Kestila, H. Jalanko, C. Holmberg, and K. Tryggvason. 1999. Nephrin is specifically located at the slit diaphragm of glomerular podocytes. *Proceedings of the National Academy of Sciences of the United States of America*. 96:7962-7.

Rupp, F., D.G. Payan, C. Magill-Solc, D.M. Cowan, and R.H. Scheller. 1991. Structure and expression of a rat agrin. *Neuron*. 6:811-23.

Saleem, M.A., L. Ni, I. Witherden, K. Tryggvason, V. Ruotsalainen, P. Mundel, and P.W. Mathieson. 2002. Co-localization of nephrin, podocin, and the actin cytoskeleton: evidence for a role in podocyte foot process formation. *American Journal of Pathology*. 161:1459-66.

Saus, J., S. Quinones, A. MacKrell, B. Blumberg, G. Muthukumaran, T. Pihlajaniemi, and M. Kurkinen. 1989. The complete primary structure of mouse alpha 2(IV) collagen. Alignment with mouse alpha 1(IV) collagen. *Journal of Biological Chemistry*. 264:6318-24.

Saus, J., J. Wieslander, J.P. Langeveld, S. Quinones, and B.G. Hudson. 1988. Identification of the Goodpasture antigen as the alpha 3(IV) chain of collagen IV. *Journal of Biological Chemistry*. 263:13374-80.

Saxena, R., P. Bygren, R. Butkowski, and J. Wieslander. 1990. Entactin: a possible auto-antigen in the pathogenesis of non-Goodpasture anti-GBM nephritis. *Kidney International*. 38:263-72.

Shih, N.Y., J. Li, R. Cotran, P. Mundel, J.H. Miner, and A.S. Shaw. 2001. CD2AP localizes to the slit diaphragm and binds to nephrin via a novel C-terminal domain. *American Journal of Pathology*. 159:2303-8.

Shih, N.Y., J. Li, V. Karpitskii, A. Nguyen, M.L. Dustin, O. Kanagawa, J.H. Miner, and A.S. Shaw. 1999. Congenital nephrotic syndrome in mice lacking CD2-associated protein.[comment]. *Science*. 286:312-5.

Siebold, B., R.A. Qian, R.W. Glanville, H. Hofmann, R. Deutzmann, and K. Kuhn. 1987. Construction of a model for the aggregation and cross-linking region (7S domain) of type IV collagen based upon an evaluation of the primary structure of

the alpha 1 and alpha 2 chains in this region. *European Journal of Biochemistry*. 168:569-75.

Soininen, R., T. Haka-Risku, D.J. Prockop, and K. Tryggvason. 1987. Complete primary structure of the alpha 1-chain of human basement membrane (type IV) collagen. *FEBS Letters*. 225:188-94.

Soininen, R., M. Huotari, A. Ganguly, D.J. Prockop, and K. Tryggvason. 1989. Structural organization of the gene for the alpha 1 chain of human type IV collagen. *Journal of Biological Chemistry*. 264:13565-71.

Soininen, R., M. Huotari, S.L. Hostikka, D.J. Prockop, and K. Tryggvason. 1988. The structural genes for alpha 1 and alpha 2 chains of human type IV collagen are divergently encoded on opposite DNA strands and have an overlapping promoter region. *Journal of Biological Chemistry*. 263:17217-20.

Spear, G.S., and R.J. Slusser. 1972. Alport's syndrome. Emphasizing electron microscopic studies of the glomerulus. *American Journal of Pathology*. 69:213-24.

Spiro, R.G. 1967a. The structure of the disaccharide unit of the renal glomerular basement membrane. *Journal of Biological Chemistry*. 242:4813-23.

Spiro, R.G. 1967b. Studies on the renal glomerular basement membrane. Nature of the carbohydrate units and their attachment to the peptide portion. *Journal of Biological Chemistry*. 242:1923-32.

Steward, A.P. 1984. Familial nephropathy in the cocker spaniel. *Journal of Small Animal Practice*. 25:15-24.

Sugimoto, M., T. Oohashi, and Y. Ninomiya. 1994. The genes COL4A5 and COL4A6, coding for basement membrane collagen chains alpha 5(IV) and alpha 6(IV), are located head-to-head in close proximity on human chromosome Xq22 and COL4A6 is transcribed from two alternative promoters. *Proceedings of the National Academy of Sciences of the United States of America*. 91:11679-83.

Swarm, R.H. 1964. Transplantation of a murine chondrosarcoma in mice of different inbred strains. *J. Natl. Cancer Inst.* 31:973-978.

Szpiro-Tapia, S., G. Bobrie, M. Guilloud-Bataille, S. Heuertz, C. Julier, J. Frezal, J.P. Grunfeld, and M.C. Hors-Cayla. 1988. Linkage studies in X-linked Alport's syndrome. *Human Genetics*. 81:85-7.

Teekhasaenee, C., S. Nimmanit, S. Wutthiphan, K. Vareesangthip, T. Laohapand, P. Malasitr, and R. Ritch. 1991. Posterior polymorphous dystrophy and Alport syndrome. *Ophthalmology*. 98:1207-15.

Thorner, P., R. Baumal, V.E. Valli, D. Mahuran, R. McInnes, and P. Marrano. 1989. Abnormalities in the NC1 domain of collagen type IV in GBM in canine hereditary nephritis. *Kidney International*. 35:843-50.

Thorner, P., B. Jansen, R. Baumal, V.E. Valli, and A. Goldberger. 1987. Samoyed hereditary glomerulopathy. Immunohistochemical staining of basement membranes of kidney for laminin, collagen type IV, fibronectin, and Goodpasture antigen, and correlation with electron microscopy of glomerular capillary basement membranes. *Laboratory Investigation*. 56:435-43.

Thorner, P.S., K. Zheng, R. Kalluri, R. Jacobs, and B.G. Hudson. 1996. Coordinate gene expression of the alpha3, alpha4, and alpha5 chains of collagen type IV. Evidence from a canine model of X-linked nephritis with a COL4A5 gene mutation. *Journal of Biological Chemistry*. 271:13821-8.

Timpl, R., M. Dziadek, S. Fujiwara, H. Nowack, and G. Wick. 1983. Nidogen: a new, self-aggregating basement membrane protein. *European Journal of Biochemistry*. 137:455-65.

Timpl, R., H. Rohde, P.G. Robey, S.I. Rennard, J.M. Foidart, and G.R. Martin. 1979. Laminin--a glycoprotein from basement membranes. *Journal of Biological Chemistry*. 254:9933-7.

Timpl, R., H. Wiedemann, V. van Delden, H. Furthmayr, and K. Kuhn. 1981. A network model for the organization of type IV collagen molecules in basement membranes. *European Journal of Biochemistry*. 120:203-11.

Trueb, B., B. Grobli, M. Spiess, B.F. Odermatt, and K.H. Winterhalter. 1982. Basement membrane (type IV) collagen is a heteropolymer. *Journal of Biological Chemistry*. 257:5239-45.

Tsen, G., W. Halfter, S. Kroger, and G.J. Cole. 1995. Agrin is a heparan sulfate proteoglycan. *Journal of Biological Chemistry*. 270:3392-9.

Tsilibary, E.C., and A.S. Charonis. 1986. The role of the main noncollagenous domain (NC1) in type IV collagen self-assembly. *Journal of Cell Biology*. 103:2467-73.

Uitto, J., L.W. Murray, B. Blumberg, and A. Shamban. 1986. UCLA conference. Biochemistry of collagen in diseases. *Annals of Internal Medicine*. 105:740-56.

van den Born, J., L.P. van den Heuvel, M.A. Bakker, J.H. Veerkamp, K.J. Assmann, and J.H. Berden. 1992. A monoclonal antibody against GBM heparan sulfate induces an acute selective proteinuria in rats. *Kidney International*. 41:115-23.

van der Loop, F.T., L. Heidet, E.D. Timmer, B.J. van den Bosch, A. Leinonen, C. Antignac, J.A. Jefferson, A.P. Maxwell, L.A. Monnens, C.H. Schroder, and H.J. Smeets. 2000. Autosomal dominant Alport syndrome caused by a COL4A3 splice site mutation. *Kidney International*. 58:1870-5.

van der Rest, M., and R. Garrone. 1991. Collagen family of proteins. *FASEB Journal*. 5:2814-23.

Vuorio, E., and B. de Crombrugghe. 1990. The family of collagen genes. *Annual Review of Biochemistry*. 59:837-72.

Weber, M. 1993. Rapidly progressive glomerulonephritis: recent advances in pathogenesis, diagnosis, and therapy. *Clinical Investigator*. 71:825-9.

Wester, D.C., C.L. Atkin, and M.C. Gregory. 1995. Alport syndrome: clinical update. *Journal of the American Academy of Audiology*. 6:73-9.

Wieslander, J., J.F. Barr, R.J. Butkowski, S.J. Edwards, P. Bygren, D. Heinegard, and B.G. Hudson. 1984. Goodpasture antigen of the glomerular basement membrane: localization to noncollagenous regions of type IV collagen. *Proceedings of the National Academy of Sciences of the United States of America*. 81:3838-42.

Wong, M.A., S. Cui, and S.E. Quaggin. 2000. Identification and characterization of a glomerular-specific promoter from the human nephrin gene. *American Journal of Physiology - Renal Fluid & Electrolyte Physiology*. 279:F1027-32.

Wu, R.R., and J.R. Couchman. 1997. cDNA cloning of the basement membrane chondroitin sulfate proteoglycan core protein, bamacan: a five domain structure including coiled-coil motifs. *Journal of Cell Biology*. 136:433-44.

Wu, T.C., Y.J. Wan, A.E. Chung, and I. Damjanov. 1983. Immunohistochemical localization of entactin and laminin in mouse embryos and fetuses. *Developmental Biology*. 100:496-505.

Yoshioka, K., S. Hino, T. Takemura, S. Maki, J. Wieslander, Y. Takekoshi, H. Makino, M. Kagawa, Y. Sado, and C.E. Kashtan. 1994. Type IV collagen alpha 5 chain. Normal distribution and abnormalities in X-linked Alport syndrome revealed by monoclonal antibody. *American Journal of Pathology*. 144:986-96.

Yuan, H., E. Takeuchi, G.A. Taylor, M. McLaughlin, D. Brown, and D.J. Salant. 2002. Nephrin dissociates from actin, and its expression is reduced in early experimental membranous nephropathy. *Journal of the American Society of Nephrology*. 13:946-56.

Yurchenco, P.D., Y.S. Cheng, and H. Colognato. 1992. Laminin forms an independent network in basement membranes.[erratum appears in J Cell Biol 1992 Jun;118(2):493]. *Journal of Cell Biology*. 117:1119-33.

Yurchenco, P.D., and H. Furthmayr. 1984. Self-assembly of basement membrane collagen. *Biochemistry*. 23:1839-50.

Yurchenco, P.D., and G.C. Ruben. 1987. Basement membrane structure in situ: evidence for lateral associations in the type IV collagen network. *Journal of Cell Biology*. 105:2559-68.

Yurchenco, P.D., E.C. Tsilibary, A.S. Charonis, and H. Furthmayr. 1985. Laminin polymerization in vitro. Evidence for a two-step assembly with domain specificity. *Journal of Biological Chemistry*. 260:7636-44.

Zheng, K., P.S. Thorner, P. Marrano, R. Baumal, and R.R. McInnes. 1994. Canine X chromosome-linked hereditary nephritis: a genetic model for human X-linked hereditary nephritis resulting from a single base mutation in the gene encoding the alpha 5 chain of collagen type IV. *Proceedings of the National Academy of Sciences of the United States of America*. 91:3989-93.

Zhou, J., M. Ding, Z. Zhao, and S.T. Reeders. 1994a. Complete primary structure of the sixth chain of human basement membrane collagen, alpha 6(IV). Isolation of the cDNAs for alpha 6(IV) and comparison with five other type IV collagen chains. *Journal of Biological Chemistry*. 269:13193-9.

Zhou, J., J.M. Hertz, A. Leinonen, and K. Tryggvason. 1992. Complete amino acid sequence of the human alpha 5 (IV) collagen chain and identification of a single-base mutation in exon 23 converting glycine 521 in the collagenous domain to cysteine in an Alport syndrome patient. *Journal of Biological Chemistry*. 267:12475-81.

Zhou, J., A. Leinonen, and K. Tryggvason. 1994b. Structure of the human type IV collagen COL4A5 gene. *Journal of Biological Chemistry*. 269:6608-14.

APPENDIX 1

Canine COL4A3 nucleotide sequence

Sequence Range: 1 to 4643

```
          10          20          30          40          50
AACTT TAAGA AGGAG ATATA CCATG GCGAT GGTGT ACCAG GATGC AATGG
TTGAA ATTCT TCCTC TATAT GGTAC CGCTA CCACA TGGTC CTACG TTACC

          60          70          80          90         100
TTCTA AGGGT GAGCA AGGGT TTCCA GGCCT TCCAG GGACA CCAGG CTACC
AAGAT TCCCA CTCGT TCCCA AAGGT CCGGA AGGTC CCTGT GGTCC GATGG

         110         120         130         140         150
CAGGG ATCCC GGGTG CTGTT GGCCT GAAAG GAGAA AAAGG TGCTC CTGCC
GTCCC TAGGG CCCAC GACAA CCGGA CTTTC CTCTT TTTCC ACGAG GACGG

         160         170         180         190         200
ACAGA AGGTA TAGAA CTTGA TGGAA GAGGT GACCC TGGGT TGCCA GGAGC
TGTCT TCCAT ATCTT GAACT ACCTT CTCCA CTGGG ACCCA ACGGT CCTCG

         210         220         230         240         250
TCCAG GATTC CAGGG TTTAC CAGGC CTTCC AGGCT TTCCG GGACC TGCTG
AGGTC CTAAG GTCCC AAATG GTCCG GAAGG TCCGA AAGGC CCTGG ACGAC

         260         270         280         290         300
GTCCA CCTGG CCCTC CGGGA TTCTT AGGCT TTCCA GGAGC CATGG GACCT
CAGGT GGACC GGGAG GCCCT AAGAA TCCGA AAGGT CCTCG GTACC CTGGA

         310         320         330         340         350
CCAGG ACCTA AGGGT CAAAT GGGCG ATAAA GTGAT AGGAC AAAAA GGAGA
GGTCC TGGAT TCCCA GTTTA CCCGC TATTT CACTA TCCTG TTTTT CCTCT

         360         370         380         390         400
GCGGG GTGTG AAAGG ATTAA CAGGA CCTCC CGGAC CACCA GGAAC AGTTA
CGCCC CACAC TTTCC TAATT GTCCT GGAGG GCCTG GTGGT CCTTG TCAAT

         410         420         430         440         450
TTGTG ACGCT AACCG GCCCT GATAA CAGAA CGGAC CTCAA GGGGG AGAAG
AACAC TGCGA TTGGC CGGGA CTATT GTCTT GCCTG GAGTT CCCCC TCTTC

         460         470         480         490         500
GGAGA CAAGG GAGCC ATGGG GAAGC CTGGA CCTCC TGGGC CCTTG GGACC
CCTCT GTTCC CTCGG TACCC CTTCG GACCT GGAGG ACCCG GGAAC CCTGG

         510         520         530         540         550
GCCTG GAGAA TCCTA TGGAT CTGAA AAAGG TGCTC CTGGA GAACC TGGCC
CGGAC CTCTT AGGAT ACCTA GACTT TTTCC ACGAG GACCT CTTGG ACCGG

         560         570         580         590         600
CGCAG GGAAA ACCTG GCAAA GATGG TGCCC CTGGT TTCCC TGGCA CTGAG
GCGTC CCTTT TGGAC CGTTT CTACC ACGGG GACCA AAGGG ACCGT GACTC
```

```
        610         620         630         640         650
GGGCC CAAAG GCAAC AGAGG CTTCC CTGGG TTACG GGGTG AAGAC GGCAT
CCCGG GTTTC CGTTG TCTCC GAAGG GACCC AATGC CCCAC TTCTG CCGTA

        660         670         680         690         700
TAAGG GGTGG AAAGG GGACA TTGGC CCGCC AGGAT TTCGT GGTCC AACAG
ATTCC CCACC TTTCC CCTGT AACCG GGCGG TCCTA AAGCA CCAGG TTGTC

        710         720         730         740         750
AATAT TATGA TGCAT ACCAG GAAAA GGGTG ATGAA GGAAT TCCAG GCCCA
TTATA ATACT ACGTA TGGTC CTTTT CCCAC TACTT CCTTA AGGTC CGGGT

        760         770         780         790         800
CCAGG CCCCA AAGGA GCTCG TGGCC CCCAG GGTCC CAGCG GTCCC CCTGG
GGTCC GGGGT TTCCT CGAGC ACCGG GGGTC CCAGG GTCGC CAGGG GGACC

        810         820         830         840         850
AGTTC TTGGA AGTGC TGGGT CATCA AAGCC TGGCC TCAGA GGAGC TCCTG
TCAAG AACCT TCACG ACCCA GTAGT TTCGG ACCGG AGTCT CCTCG AGGAC

        860         870         880         890         900
GATCT CCAGG CATGA AAGGA CGTAA AGGGG AACAA GGGCC CCCGG GAAAG
CTAGA GGTCC GTACT TTCCT GCATT TCCCC TTGTT CCCGG GGGCC CTTTC

        910         920         930         940         950
AATGC AGTGG GGCCT CCTGG GTCCC CAGGT TGTCC TGGTT CACCA GGCCC
TTACG TCACC CCGGA GGACC CAGGG GTCCA ACAGG ACCAA GTGGT CCGGG

        960         970         980         990        1000
CATAG GGCTG CCGGG ATATC CAGGA CCACC AGGTG ACATC GTGTT TCGCA
GTATC CCGAC GGCCC TATAG GTCCT GGTGG TCCAC TGTAG CACAA AGCGT

       1010        1020        1030        1040        1050
AAGGT CCACC TGGGG TTGGT GGACT CCCCG GCCAT GTTGG GTTTC CAGGC
TTCCA GGTGG ACCCC AACCA CCTGA GGGGC CGGTA CAACC CAAAG GTCCG

       1060        1070        1080        1090        1100
ATCCC AGGAG TCGAG GGGCC CAAAG GGGAA CCAGG CCTCT TGTGC GTACA
TAGGG TCCTC AGCTC CCCGG GTTTC CCCTT GGTCC GGAGA ACACG CATGT

       1110        1120        1130        1140        1150
GTGTG CTTAT ATCCC AGGGC GTCCA GGTCT CCCAG GACTG CCAGG GTTGG
CACAC GAATA TAGGG TCCCG CAGGT CCAGA GGGTC CTGAC GGTCC CAACC

       1160        1170        1180        1190        1200
ATGGC ATAAA AGGAA TCCCA GGAGG ACAAG GGGCA GCTGG CATTA AAGGA
TACCG TATTT TCCTT AGGGT CCTCC TGTTC CCCGT CGACC GTAAT TTCCT

       1210        1220        1230        1240        1250
AGCCC AGGGT CCCCA GGAAG TGCGG GTCTT CCAGG ATTTC CAGGA TTCCC
TCGGG TCCCA GGGGT CCTTC ACGCC CAGAA GGTCC TAAAG GTCCT AAGGG

       1260        1270        1280        1290        1300
GGGTG CTCCG GGTGC TCCAG GACTT AAAGG AGAAA AAGGA GAAAC ATCTG
CCCAC GAGGC CCACG AGGTC CTGAA TTTCC TCTTT TTCCT CTTTG TAGAC
```

```
       1310        1320        1330        1340        1350
AGCCT GAGGG ACAGG TGGGT GCTCC AGGGG ACCCC GGGCT CAGAG GCCAT
TCGGA CTCCC TGTCC ACCCA CGAGG TCCCC TGGGG CCCGA GTCTC CGGTA

       1360        1370        1380        1390        1400
CCTGG AAGAA AGGGC TTGGA TGGAA TTCCT GGAAC TCCTG GAATT AAAGG
GGACC TTCTT TCCCG AACCT ACCTT AAGGA CCTTG AGGAC CTTAA TTTCC

       1410        1420        1430        1440        1450
ACCAC CAGGA CCCAA AGGTG AACCG GCCCT GAGTG GTGAG AAGGG GGACC
TGGTG GTCCT GGGTT TCCAC TTGGC CGGGA CTCAC CACTC TTCCC CCTGG

       1460        1470        1480        1490        1500
AGGGT CCTCC AGGGG ATCCT GGGAT CCCAG GGTCC CCAGG ACCTG CAGGA
TCCCA GGAGG TCCCC TAGGA CCCTA GGGTC CCAGG GGTCC TGGAC GTCCT

       1510        1520        1530        1540        1550
CCACC TGGAC CACCA AACTA TGGAC CACAG GGAGA GCCTG GTCCA AAGGG
GGTGG ACCTG GTGGT TTGAT ACCTG GTGTC CCTCT CGGAC CAGGT TTCCC

       1560        1570        1580        1590        1600
CACCC AAGGA GTTCC TGGAG CCCCT GGACC ACCTG GAGAA GCCGG TCCTA
GTGGG TTCCT CAAGG ACCTC GGGGA CCTGG TGGAC CTCTT CGGCC AGGAT

       1610        1620        1630        1640        1650
AGGGA GAATT CGGTA TTTCG ACACC AGTCC CAGGG CCCCC AGGAC CTCCA
TCCCT CTTAA GCCAT AAAGC TGTGG TCAGG GTCCC GGGGG TCCTG GAGGT

       1660        1670        1680        1690        1700
GGGCC CCCTG GTTAT GCTGG CCCCC GAGGT CCACC TGGTA TCCCT GGATC
CCCGG GGGAC CAATA CGACC GGGGG CTCCA GGTGG ACCAT AGGGA CCTAG

       1710        1720        1730        1740        1750
CATAG GAAAA TGTGG TGATC CGGGT CTTCC TGGGC CTGAT GGTGA ACCAG
GTATC CTTTT ACACC ACTAG GCCCA GAAGG ACCCG GACTA CCACT TGGTC

       1760        1770        1780        1790        1800
GAATT CCAGG AATTG GCTTC CCTGG GCCCC CTGGA CCTAA GGGAG ACCAA
CTTAA GGTCC TTAAC CGAAG GGACC CGGGG GACCT GGATT CCCTC TGGTT

       1810        1820        1830        1840        1850
GGTTT TCCAG GAGCA AAAGG AGCAC CGGGT TGTCC AGGAG AAATG GGGAA
CCAAA AGGTC CTCGT TTTCC TCGTG GCCCA ACAGG TCCTC TTTAC CCCTT

       1860        1870        1880        1890        1900
GCCCG GGTCA CCTGG AGAAC CAGGT CTCCC AGGAG CCAAG GGAGA ACCAG
CGGGC CCAGT GGACC TCTTG GTCCA GAGGG TCCTC GGTTC CCTCT TGGTC

       1910        1920        1930        1940        1950
GACTA GCCAT ACCTG GAGAA CCAGG AATAC CAGGT TTTCC AGGAG AAAGA
CTGAT CGGTA TGGAC CTCTT GGTCC TTATG GTCCA AAAGG TCCTC TTTCT

       1960        1970        1980        1990        2000
GGCAA TTCTG GGGAA AATGG AGAAA TTGGA CTCCC TGGAC TTCCA GGTCT
CCGTT AAGAC CCCTT TTACC TCTTT AACCT GAGGG ACCTG AAGGT CCAGA
```

```
       2010        2020        2030        2040        2050
CCCTG GAATT CCAGG AACTG GAGGG CTTGA TGGAC CACCA GGGGA TCCAG
GGGAC CTTAA GGTCC TTGAC CTCCC GAACT ACCTG GTGGT CCCCT AGGTC

       2060        2070        2080        2090        2100
GGAAG CCTGG ACCAC CTGGA GAAAA AGGAC CCCCA GGAAG TTGCA CAGAG
CCTTC GGACC TGGTG GACCT CTTTT TCCTG GGGGT CCTTC AACGT GTCTC

       2110        2120        2130        2140        2150
GGTCC CAGAG GAGCC CACGG ACTTC CAGGA TTAAA TGGAT TGGAA GGGCA
CCAGG GTCTC CTCGG GTGCC TGAAG GTCCT AATTT ACCTA ACCTT CCCGT

       2160        2170        2180        2190        2200
ACCAG GCAGA AGAGG TGAAA CAGGG CCAAA AGGAG ACCCA GGTAT TCCAG
TGGTC CGTCT TCTCC ACTTT GTCCC GGTTT TCCTC TGGGT CCATA AGGTC

       2210        2220        2230        2240        2250
GCTTG GATAG ATCAG GCTTT CCTGG AGAAC CTGGA CCACC AGGAA TGCCA
CGAAC CTATC TAGTC CGAAA GGACC TCTTG GACCT GGTGG TCCTT ACGGT

       2260        2270        2280        2290        2300
GGTCA TCGAG GTGAG ATAGG ACCAC CTGGC CAAAA AGGAT ATCCA GGAAA
CCAGT AGCTC CACTC TATCC TGGTG GACCG GTTTT TCCTA TAGGT CCTTT

       2310        2320        2330        2340        2350
TCCGG GATTT TTAGG ACTAC CAGGT GAAGA TGGAG TGATT GGGAT GATGG
AGGCC CTAAA AATCC TGATG GTCCA CTTCT ACCTC ACTAA CCCTA CTACC

       2360        2370        2380        2390        2400
GCTAT CCAGG AAACA CTGGC CCTCC TGGGC CTCCC GGGAT CCCAG GCACC
CGATA GGTCC TTTGT GACCG GGAGG ACCCG GAGGG CCCTA GGGTC CGTGG

       2410        2420        2430        2440        2450
CCAGG ACAGA GGGGT AGCTT TGGAA TTCCA GGTGC AAAGG GTGAG AAAGG
GGTCC TGTCT CCCCA TCGAA ACCTT AAGGT CCACG TTTCC CACTC TTTCC

       2460        2470        2480        2490        2500
GCCCC CAGGA GCCAA GGGGG AAGAA GGAGA GAAAG GACCT ATGGG GCCTT
CGGGG GTCCT CGGTT CCCCC TTCTT CCTCT CTTTC CTGGA TACCC CGGAA

       2510        2520        2530        2540        2550
GTCAA ATAAC CAGCT CAGTG GGGGA CAAAG GAGAA CCAGG CCTCA AAGGG
CAGTT TATTG GTCGA GTCAC CCCCT GTTTC CTCTT GGTCC GGAGT TTCCC

       2560        2570        2580        2590        2600
TTTGC AGGAA AGCCG GGTGA AAAAG GAAAC AGAGG CATTC CAGGG TTACC
AAACG TCCTT TCGGC CCACT TTTTC CTTTG TCTCC GTAAG GTCCC AATGG

       2610        2620        2630        2640        2650
AGGTT TGAAA GGATT CGAAG GGCCA CCTGG ACCAC CAGGC CCACC AGGCC
TCCAA ACTTT CCTAA GCTTC CCGGT GGACC TGGTG GTCCG GGTGG TCCGG

       2660        2670        2680        2690        2700
CCCGA GGAGA TCCAG GCAGC ATTGG GAATC CTGGA GAAGC AGGAC CACGT
GGGCT CCTCT AGGTC CGTCG TAACC CTTAG GACCT CTTCG TCCTG GTGCA
```

```
       2710        2720        2730        2740        2750
GGAAG TCCAG GAAAC ATGGG GAACA TGGGG ATGCC AGGTC TTAAA GGACA
CCTTC AGGTC CTTTG TACCC CTTGT ACCCC TACGG TCCAG AATTT CCTGT

       2760        2770        2780        2790        2800
CAGGG GACCT TTGGG ACTAC CAGGT GTCAC TGGAA GACCA GGCCT CCCAG
GTCCC CTGGA AACCC TGATG GTCCA CAGTG ACCTT CTGGT CCGGA GGGTC

       2810        2820        2830        2840        2850
GTGTT CACGG TCTCC AAGGA GACAA GGGAG AGCCA GGTTA TTCAG CAGGT
CACAA GTGCC AGAGG TTCCT CTGTT CCCTC TCGGT CCAAT AAGTC GTCCA

       2860        2870        2880        2890        2900
ACAAG GCCAG GACCA CCGGG ACTAA AGGGA GATCC AGGAT TACCA GGTGC
TGTTC CGGTC CTGGT GGCCC TGATT TCCCT CTAGG TCCTA ATGGT CCACG

       2910        2920        2930        2940        2950
CATGG GAAGG AAAGG AGAAA GAGGG CTACC TGGCA CCCCT GGCCA TTCGG
GTACC CTTCC TTTCC TCTTT CTCCC GATGG ACCGT GGGGA CCGGT AAGCC

       2960        2970        2980        2990        3000
GGCCT GCTGG AACTG AGGGA GCCCC TGGAC ATCCC GGAAG TCCTG GCCAC
CCGGA CGACC TTGAC TCCCT CGGGG ACCTG TAGGG CCTTC AGGAC CGGTG

       3010        3020        3030        3040        3050
CCAGG AAAGC CAGGC CCTGA TGGAG ATTTG GGGTT AAAAG GCATC AAAGG
GGTCC TTTCG GTCCG GGACT ACCTC TAAAC CCCAA TTTTC CGTAG TTTCC

       3060        3070        3080        3090        3100
CTTCC CTGGG TCCCC AGGAG TCAAA GGACC TCCAG GACCT CCAGG AATCC
GAAGG GACCC AGGGG TCCTC AGTTT CCTGG AGGTC CTGGA GGTCC TTAGG

       3110        3120        3130        3140        3150
TAGGA TCTCC TGGAC CTATG GGCAT GAGAG GCAAC CAGGG CCGTG ATGGA
ATCCT AGAGG ACCTG GATAC CCGTA CTCTC CGTTG GTCCC GGCAC TACCT

       3160        3170        3180        3190        3200
ATCCC TGGTC CCCCA GGAGA AAAGG GAGAA ACAGG TTTGC TGGGG GCACA
TAGGG ACCAG GGGGT CCTCT TTTCC CTCTT TGTCC AAACG ACCCC CGTGT

       3210        3220        3230        3240        3250
TCCAG GCCCA GGAGG GAGCC CTGGT CCTCC AGGAG CCAAA GGAGA CAGGG
AGGTC CGGGT CCTCC CTCGG GACCA GGAGG TCCTC GGTTT CCTCT GTCCC

       3260        3270        3280        3290        3300
GCGCC CCAGG CCTAC CAGGC CCCCC AGGCA GGAAA GGGGC AGTGG GAGAT
CGCGG GGTCC GGATG GTCCG GGGGG TCCGT CCTTT CCCCG TCACC CTCTA

       3310        3320        3330        3340        3350
GCCGG GCCAC GGGGA CCCAC TGGCA TGACG GGACT CCCAG GGCCA CCAGG
CGGCC CGGTG CCCCT GGGTG ACCGT ACTGC CCTGA GGGTC CCGGT GGTCC

       3360        3370        3380        3390        3400
TTTTC CTGGT GCAGT CATCC CTGGC CAGAA AGGAA ATCAA GGTCC ACCAG
AAAAG GACCA CGTCA GTAGG GACCG GTCTT TCCTT TAGTT CCAGG TGGTC
```

```
       3410        3420        3430        3440        3450
GCTTC AGAGG AAACC CAGGT GAGCC TGGTC CTCCA GGCCC TCCAG GGAGC
CGAAG TCTCC TTTGG GTCCA CTCGG ACCAG GAGGT CCGGG AGGTC CCTCG

       3460        3470        3480        3490        3500
CACGT GAAAG GCATA AAAGG AGACA AGGGA CTCAT GGGTG AGCCT GGCCC
GTGCA CTTTC CGTAT TTTCC TCTGT TCCCT GAGTA CCCAC TCGGA CCGGG

       3510        3520        3530        3540        3550
CAGAG GTCTA CCTGG AACTG TGGGA GACAA GGGGC CACCG GGTCA ACCGG
GTCTC CAGAT GGACC TTGAC ACCCT CTGTT CCCCG GTGGC CCAGT TGGCC

       3560        3570        3580        3590        3600
GAGCA CCAGG GAGCC CAGGT CTGCC AGGGC TCAGG GGCGA TCCTG GATTC
CTCGT GGTCC CTCGG GTCCA GACGG TCCCG AGTCC CCGCT AGGAC CTAAG

       3610        3620        3630        3640        3650
TATGG ATTTC CAGGT GTTAA AGGAG AGAAG GGTAA CCCAG GATTT CCGGG
ATACC TAAAG GTCCA CAATT TCCTC TCTTC CCATT GGGTC CTAAA GGCCC

       3660        3670        3680        3690        3700
ACCAG CTGGA CCTCC AGGGC AAATT GGGCC AAAAG GACCA CCTGG TGTCC
TGGTC GACCT GGAGG TCCCG TTTAA CCCGG TTTTC CTGGT GGACC ACAGG

       3710        3720        3730        3740        3750
GTGGA GACCC TGGCA CAGTT AAGAT CATCT CCCTT CCAGG AAGCC CAGGG
CACCT CTGGG ACCGT GTCAA TTCTA GTAGA GGGAA GGTCC TTCGG GTCCC

       3760        3770        3780        3790        3800
CCACC TGGCC GTGCT GGAGG ACCGG GGATG CCAGG AGAAC CAGGG CCACC
GGTGG ACCGG CACGA CCTCC TGGCC CCTAC GGTCC TCTTG GTCCC GGTGG

       3810        3820        3830        3840        3850
AGGGC CACCA GGAAT CCTAG GACCC TGTGG ACCAA GAGGT AAACC AGGCA
TCCCG GTGGT CCTTA GGATC CTGGG ACACC TGGTT CTCCA TTTGG TCCGT

       3860        3870        3880        3890        3900
TGGAT GGAAT ACCAG GAACT CCTGG GCCAA TTGGA GAAAA AGGCA ACAAA
ACCTA CCTTA TGGTC CTTGA GGACC CGGTT AACCT CTTTT TCCGT TGTTT

       3910        3920        3930        3940        3950
GGTTG TAAAG GCGAG CAAGG ACCAC CTGGA TTAGA TGGAC TGCCA GGCTT
CCAAC ATTTC CGCTC GTTCC TGGTG GACCT AATCT ACCTG ACGGT CCGAA

       3960        3970        3980        3990        4000
GAAGG GGAGA CCTGG AGACC CTGGA CCACC TGCAA CTGGG ACCAT GAAGA
CTTCC CCTCT GGACC TCTGG GACCT GGTGG ACGTT GACCC TGGTA CTTCT

       4010        4020        4030        4040        4050
GGGGC TTTAT CTTCA CCAGG CATAG TCAGA CCACA GTGAT TCCTT CCTGT
CCCCG AAATA GAAGT GGTCC GTATC AGTCT GGTGT CACTA AGGAA GGACA

       4060        4070        4080        4090        4100
CCAGA AGGGA CAGCG CCACT CTATA GTGGC TTTTC TCTTC TTTTC ATACA
GGTCT TCCCT GTCGC GGTGA GATAT CACCG AAAAG AGAAG AAAAG TATGT
```

```
       4110        4120        4130        4140        4150
AGGAA ATGAA CAAGC CCATG GACAA GACTT GGGAA CTCTC GGCAG CTGCC
TCCTT TACTT GTTCG GGTAC CTGTT CTGAA CCCTT GAGAG CCGTC GACGG

       4160        4170        4180        4190        4200
TGCAG CGATT TACCA CAATG CCATT CTTGT TCTGT AACAT CAACA ATGTC
ACGTC GCTAA ATGGT GTTAC GGTAA GAACA AGACA TTGTA GTTGT TACAG

       4210        4220        4230        4240        4250
TGTAA TTTTG CATCT CGAAA TGACT ACTCA TACTG GCTGT CAACA CCAGC
ACATT AAAAC GTAGA GCTTT ACTGA TGAGT ATGAC CGACA GTTGT GGTCG

       4260        4270        4280        4290        4300
TCTGA TGCCA ATGGA CATGG CTCCA ATTAC TGGCA GGGCC CTGGA GCCTT
AGACT ACGGT TACCT GTACC GAGGT TAATG ACCGT CCCGG GACCT CGGAA

       4310        4320        4330        4340        4350
ATATT AGCAG ATGCA CTGTC TGTGA AGGTC CTACG ATTGC CATAG CCATT
TATAA TCGTC TACGT GACAG ACACT TCCAG GATGC TAACG GTATC GGTAA

       4360        4370        4380        4390        4400
CACAG CCAAA CCACT GACAT TCCCT CATGT CCCAA TGGCT GGATT TCTCT
GTGTC GGTTT GGTGA CTGTA AGGGA GTACA GGGTT ACCGA CCTAA AGAGA

       4410        4420        4430        4440        4450
CTGGA AAGGA TTTTC GTTTA TCATG TTCAC AAGTG CAGGT TCTGA GGGTG
GACCT TTCCT AAAAG CAAAT AGTAC AAGTG TTCAC GTCCA AGACT CCCAC

       4460        4470        4480        4490        4500
CTGGG CAAGC ACTGG CCTCC CCCGG CTCCT GCCTG GAAGA GTTCC GAGCC
GACCC GTTCG TGACC GGAGG GGGCC GAGGA CGGAC CTTCT CAAGG CTCGG

       4510        4520        4530        4540        4550
AGTCC GTTTA TAGAA TGTCA CGGAA GAGGA ACGTG CAACT ACTAT TCAAA
TCAGG CAAAT ATCTT ACAGT GCCTT CTCCT TGCAC GTTGA TGATA AGTTT

       4560        4570        4580        4590        4600
TTCCT ACAGT TTCTG GTTGG CTTCG TTAAA CCCCC AAAGA ATGTT CAGAA
AAGGA TGTCA AAGAC CAACC GAAGC AATTT GGGGG TTTCT TACAA GTCTT

       4610        4620        4630        4640
AACCT ATTCC ATCAA CTGTG AAAGC TGAAT CCATA TGACT AGT
TTGGA TAAGG TAGTT GACAC TTTCG ACTTA GGTAT ACTGA TCA
```

APPENDIX 2

Canine COL4A3 nucleotide alignments

```
Consensus key (see documentation for details)
* - single, fully conserved residue
  - no consensus


CLUSTAL W (1.81) multiple sequence alignment


Canine_Normal_5-18      ------------------------------------------------------------
Human_COL4A3            ATGAGCGCCCGGACCGCCCCCAGGCCGCAGGTGCTCCTGCTGCCGCTCCTGCTGGTGCTC
Mouse_COL4A3            ATGCACTCCAAGACTGCTCCAAGGTTCCTGGTGTTCCTGCTGCTTACCCTGCTGCTACTC


Canine_Normal_5-18      ------------------------------------------------------------
Human_COL4A3            CTGGCGGCGGCGCCCGCAGCCAGCAAGGGTTGTGTCTGTAAAGACAAAGGCCAGTGCTTC
Mouse_COL4A3            CTGGCTGCCTCTCCTGTGGCTAGCAAGGGCTGTGTCTGCAAAGGCAAAGGACAATGCCTC


Canine_Normal_5-18      ------------------------------------------------------------
Human_COL4A3            TGTGACGGGGCCAAAGGGGAGAAGGGGGAGAAGGGCTTTCCTGGACCCCCCGGTTCTCCT
Mouse_COL4A3            TGCGCGGGGACCAAAGGGGAGAAGGGGGAGAAAGGGGTTCCTGGTTCCCCTGGATTTCCT


Canine_Normal_5-18      ------------------------------------------------------------
Human_COL4A3            GGCCAGAAAGGATTCACAGGTCCTGAAGGCTTGCCTGGACCGCAGGGACCCAAGGGCTTT
Mouse_COL4A3            GGCCAGAAAGGATTTCCAGGTCCTGAAGGCTTGCCTGGACCACAGGGACCCAAGGGCTCC


Canine_Normal_5-18      ------------------------------------------------------------
Human_COL4A3            CCAGGACTTCCAGGACTCACGGGTTCCAAAGGTGTAAGGGGAATAAGTGGATTGCCAGGA
Mouse_COL4A3            CCAGGACTTCCGGGACTCACTGGCCCCAAAGGCATCAGGGGAATAACTGGATTACCAGGG


Canine_Normal_5-18      --------------------------AACTTTAAGAAG----GAGATATACCATGGCGAT
Human_COL4A3            TTTTCTGGTTCTCCTGGACTTCCAGGCACCCCAGGCAATACCGGGCCTTACGGACTTGTC
Mouse_COL4A3            TTTGCAGGTCCTCCTGGACTTCCAGGCCTCCCAGGCCACCCTGGGCCTCGTGGGCTGGCT
                                                  * *       * *           *

Canine_Normal_5-18      GGTGTACCAGGATGCAATGGTTCTAAGGGTGAGCAAGGGTTTCCAGGCCTTCCAGGGACA
Human_COL4A3            GGTGTACCAGGATGCAGTGGTTCTAAGGGTGAGCAGGGGTTTCCAGGACTCCCAGGGACA
Mouse_COL4A3            GGTTTACCAGGATGCAACGGATCTAAGGGTGAACAAGGATTCCCGGGCTTTCCCGGCACA
                        *** ************  ** *********** ** ** ** ** **  * ** ** ***

Canine_Normal_5-18      CCAGGCTACCCAGGGATCCCGGGTGCTGTTGGCCTGAAAGGAGAAAAAGGTGCTCCTGCC
Human_COL4A3            CCGGGCTACCCAGGGATCCCGGGTGCTGCTGGTTTGAAAGGACAAAAGGGTGCTCCTGCT
Mouse_COL4A3            CCAGGCTATGCAGGGCTCCCAGGTCCTGATGGCTTGAAAGGACAAAAGGGTGAGCCTGCT
                        ** *****  ***** **** *** *** ***  ******** **** ****  *****

Canine_Normal_5-18      ACAG---AAGGTATAGAACTTGATGGAAGAGGTGACCCTGGGTTGCCAGGAGCTCCAGGA
Human_COL4A3            AAAGGAGAAGATATAGAACTTGATGCAAAAGGCGACCCCGGGTTGCCAGGGGCTCCAGGA
Mouse_COL4A3            CAAGGAGAAGACAGGGGATTCAATGGAAAAGGTGACCCTGGGCCTCCAGGGGTTCCAGGC
                          **   ***  *  * * *  *** ** *** ***** ***   ***** * ******

Canine_Normal_5-18      TTCCAGGGTTTACCAGGCCTTCCAGGCTTTCCGGGACCTGCTGGTCCACCTGGCCCTCCG
Human_COL4A3            CCCCAGGGTTTGCCAGGCCCTCCAGGTTTTCCTGGGCCTGTTGGCCCACCTGGTCCTCCG
Mouse_COL4A3            TTCCAGGGTTTTCCGGGACTCCCAGGTTTTCCAGGGCCTGCCGGTCCACCAGGACCTCCG
                          ********* ** ** *  ***** ***** ** ****  ** ***** ** ******

Canine_Normal_5-18      GGATTCTTAGGCTTTCCAGGAGCCATGGGACCTCCAGGACCTAAGGGTCAAATGGGCGAT
Human_COL4A3            GGATTCTTTGGCTTTCCAGGAGCCATGGGACCTAGAGGACCTAAGGGTCACATGGGTGAA
Mouse_COL4A3            GGATTTTTTGGTTTACCAGGAGCAATGGGACCCAGAGGACCCAAAGGCCACATGGGCGAT
                        ***** ** ** ** ******** ********   ****** ** ** ** ***** **
```

```
Canine_Normal_5-18      AAAGTGATAGGACAAAAAGGAGAGCGGGGTGTGAAAGGATTAACAGGACCTCCCGGACCA
Human_COL4A3            AGAGTGATAGGACATAAAGGAGAGCGGGGTGTGAAAGGGTTAACAGGACCCCCGGGACCA
Mouse_COL4A3            AGCGTGATAGGACAAAAAGGAGAAAGGGGTATGAAAGGATTAACAGGACCTCCTGGGCCA
                        *  *********** ********  ***** ******* ********** ** ** ***

Canine_Normal_5-18      CCAGGAACAGTTATTGTGACGCTAACCGGCCCTGATAACAGAACGGACCTCAAGGGGGAG
Human_COL4A3            CCAGGAACAGTTATTGTGACCCTAACTGGCCCAGATAACAGAACGGACCTCAAGGGGGAA
Mouse_COL4A3            CCAGGAACAGTGATTTTTACACTCACCCAGCCATACAACAAATCGGACTTCAAGGGAGAG
                        *********** *** * ** ** **    **  * **** * ***** ******* **

Canine_Normal_5-18      AAGGGAGACAAGGGAGCCATGGGGAAGCCTGGACCTCCTGGGCCCTTGGGACCGCCTGGA
Human_COL4A3            AAGGGAGACAAGGGAGCAATGGGCGAGCCTGGACCTCCTGGACCCTCAGGACTGCCTGGA
Mouse_COL4A3            AAAGGAGATGAGGGAGAGAGGGGTGAACCCGGACCTCCTGGACCCTCGGGGCCACCTGGA
                        ** *****  ******  * ***  * ** *********** ****  ** *  ******

Canine_Normal_5-18      GAATCCTATGGATCTGAAAAAGGTGCTCCTGGAGAACCTGGCCCGCAGGGAAAACCTGGC
Human_COL4A3            GAATCATATGGATCTGAAAAGGGTGCTCCTGGAGACCCTGGCCTGCAGGGAAAACCCGGA
Mouse_COL4A3            GACTCCTATGGATCAGAAAAGGGTGCGCCTGGAGAGCCTGGTCCTCGGGGCAAACCTGGA
                        ** ** ******** ***** ***** ******** ***** *  * *** ***** **

Canine_Normal_5-18      AAAGATGGTGCCCCTGGTTTCCCTGGCACTGAGGGGCCCAAAGGCAACAGAGGCTTCCCT
Human_COL4A3            AAAGATGGTGTTCCTGGCTTCCCTGGAAGTGAGGGAGTCAAGGGCAACAGGGGTTTCCCT
Mouse_COL4A3            AAAGATGGTGCCCCTGGCTTCCCTGGCACTGAGGGAGCCAAAGGCAACAGGGGCTTTCCT
                        **********  ***** ******** * ******   *** ******** ** ** ***

Canine_Normal_5-18      GGGTTACGGGGTGAAGACGGCATTAAGGGGTGGAAAGGGGACATTGGCCCGCCAGGATTT
Human_COL4A3            GGGTTAATGGGTGAAGATGGCATTAAGGGACAGAAAGGGGACATTGGCCCTCCAGGATTT
Mouse_COL4A3            GGGTTGAGGGGTGAAGCTGGCATTAAGGGAAGGAAAGGAGACATTGGCCCCCCAGGATTT
                        *****   ********  ***********   ****** *********** *********

Canine_Normal_5-18      CGTGGTCCAACAGAATATTATGATGCATACCAGGAAAAGGGTGATGAAGGAATTCCAGGC
Human_COL4A3            CGTGGTCCAACAGAATATTATGACACATACCAGGAAAAGGGAGATGAAGGCACTCCAGGC
Mouse_COL4A3            CCTGGTCCAACAGAATATTATGACGCATACCTGGAAAAGGGAGAGAGAGGAATGCCAGGC
                        * ********************  ****** ********* **   *** *  ******

Canine_Normal_5-18      CCACCAGGCCCCAAAGGAGCTCGTGGCCCCCAGGGTCCCAGCGGTCCCCCTGGAGTTCTT
Human_COL4A3            CCACCAGGGCCCAGAGGAGCTCGTGGCCCACAAGGTCCCAGTGGTCCCCCCGGAGTTCCT
Mouse_COL4A3            CTTCCTGGGCCCAAAGGAGCCCGTGGCCCACAGGGTCCAAGTGGTCCCCCTGGAGTCCCT
                        *  ** ** **** ****** ******** ** ***** ** ******** ***** * *

Canine_Normal_5-18      GGAAGTGCTGGGTCATCAAAGCCTGGCCTCAGAGGAGCTCCTGGATCTCCAGGCATGAAA
Human_COL4A3            GGAAGTCCTGGATCATCAAGGCCTGGCCTCAGAGGAGCCCCTGGATGGCCAGGCTTGAAA
Mouse_COL4A3            GGAAGTCCTGGACTGTCAAGGCCTGGCCTCAGAGGACCCATTGGATGGCCAGGCTTGAAA
                        ****** ****    **** ***************** *   *****  ****** *****

Canine_Normal_5-18      GGACGTAAAGGGGAACAAGGGCCCCCGGGAAAGAATGCAGTGGGGCCTCCTGGGTCCCCA
Human_COL4A3            GGAAGTAAAGGGGAACGAGGCCGCCCAGGAAAGGATGCCATGGGGACTCCTGGGTCCCCA
Mouse_COL4A3            GGGAGTAAAGGAGAGAGAGGACCCCCTGGAAAAGACACTGTGGGCCCTCCTGGACCCCTG
                        **  ******* **   *** * *** *****  *  *  ****  *******  ***

Canine_Normal_5-18      GGTTGTCCTGGTTCACCAGGCCCCATAGGGCTGCCGGGATATCCAGGACCACCAGGTGAC
Human_COL4A3            GGTTGTGCTGGTTCACCAGGTCTTCCAGGATCACCGGGACCTCCAGGACCGCCAGGTGAC
Mouse_COL4A3            GGATGTCCTGGCTCACCAGGTCCACCAGGCCCTCCAGGACCTCCAGGATGTCCAGGTGAC
                        ** *** **** ******** *    ***    ** ***  *******   *********

Canine_Normal_5-18      ATCGTGTTTCGCAAAGGTCCACCTGGGGTTGGTGGACTCCCCGGCCATGTTGGGTTTCCA
Human_COL4A3            ATCGTTTTTCGCAAGGGTCCACCTGGAGATCACGGACTGCCAGGCTATCTAGGGTCTCCA
Mouse_COL4A3            ATTGTTTTT---AAATGTTCCCCTGGTGAGCATGGAATGCCAGGTGACACAGGACCTCCA
                        ** ** ***   **  ** * ***** *     *** * ** **  *    **   ****

Canine_Normal_5-18      GGCATCCCAGGAGTCGAGGGGCCCAAAGGGGAACCAGGCCTCTTGTGCGTACAGTGTGCT
Human_COL4A3            GGAATCCCAGGAGTTGATGGGCCCAAAGGAGAACCAGGCCTCCTGTGTACACAGTGCCCT
Mouse_COL4A3            GGAGTTCCAGGACTTGATGGACCCAAAGGAGAACCAGGCAGTCCATGCACCGAGTGTCAC
                        **  * ****** * ** ** ******** *********      **     ****
```

```
Canine_Normal_5-18      TATATCCCAGGGCGTCCAGGTCTCCCAGGACTGCCAGGGTTGGATGGCATAAAAGGAATC
Human_COL4A3            TATATCCCAGGGCCTCCCGGTCTCCCAGGATTGCCAGGGTTACATGGTGTAAAAGGAATC
Mouse_COL4A3            TGCTTCCCAGGGCCCCCTGGAGTCCCAGGATTTCCAGGATTAGATGGTATCAAAGGAATC
                        *   *********  ** **  ******** * ***** **  ****  * *********

Canine_Normal_5-18      CCAGGAGGACAAGGGGCAGCTGGCATTAAAGGAAGCCCAGGGTCCCCAGGAAGTGCGGGT
Human_COL4A3            CCAGGAAGACAAGGCGCAGCTGGCTTGAAAGGAAGCCCAGGGTCCCCAGGAAATACAGGT
Mouse_COL4A3            CCAGGAGGACGAGGTGTGCCTGGTCTGAAGGGAAACCCAGGGTCCCCAGGAAGTGCAGGT
                        ****** *** *** *   ****  * ** **** ***************** * * ***

Canine_Normal_5-18      CTTCCAGGATTTCCAGGATTCCCGGGTGCTCCGGGTGCTCCAGGACTTAAAGGAGAAAAA
Human_COL4A3            CTTCCAGGATTTCCAGGTTTCCCAGGTGCCCAGGGTGACCCAGGACTTAAAGGAGAAAAA
Mouse_COL4A3            CTCCCAGGATTTGCAGGATTCCCAGGTGACCAAGGGCATCCAGGACTTAAAGGGGACAAA
                        ** ********* **** ***** ****  *  **    ************** ** ***

Canine_Normal_5-18      GGAGAAACATCTGAGCCTGAGGGACAGGTGGGTGCTCCAGGGGACCCCGGGCTCAGAGGC
Human_COL4A3            GGTGAAACACTTCAGCCTGAGGGGCAAGTGGGTGTCCCAGGTGACCCGGGGCTCAGAGGC
Mouse_COL4A3            GGGGATACACCTCTACCCTGGGGGCAAGTGGGTAATCCAGGTGATCCTGGACTCAGAGGC
                        ** ** ***  *   **   *** ** ******   ***** ** ** ** *********

Canine_Normal_5-18      CATCCTGGAAGAAAGGGCTTGGATGGAATTCCTGGAACTCCTGGAATTAAAGGACCACCA
Human_COL4A3            CAACCTGGGAGAAAGGGCTTGGATGGAATTCCTGGAACTCTGGGAGTGAAAGGATTACCA
Mouse_COL4A3            CTGCCCGGGAGAAAGGGCTTCGATGGAACTCCTGGAGGTCCAGGAGCGAAAGGACCACCA
                        *  ** ** *********** ******* *******  **  ***   ******  ****

Canine_Normal_5-18      GGACCCAAAGGTGAACCGGCCCTGAGTGGTGAGAAGGGGGACCAGGGTCCTCCAGGGGAT
Human_COL4A3            GGACCTAAAGGCGAACTGGCTCTGAGTGGTGAGAAAGGGGACCAAGGTCCTCCAGGGGAT
Mouse_COL4A3            GGACCTCAGGGCGAACCGGCCCTGAGTGGAAGGAAAGGGGACCAAGGACCTCCAGGACCT
                        *****  * ** **** *** ********   *** ******** ** ********   *

Canine_Normal_5-18      CCTGGGATCCCAGGGTCCCCAGGACCTGCAGGACCACCTGGACCACCAAACTATGGACCA
Human_COL4A3            CCTGGCTCCCCTGGGTCCCCAGGACCTGCAGGACCAGCTGGACCACCTGGCTACGGACCC
Mouse_COL4A3            CCTGGATTCCCTGGACCCCCAGGACCTGCAGGACCAGCTGGACCACCAGGCTATGGACCT
                        *****   *** **  ******************** **********   *** *****

Canine_Normal_5-18      CAGGGAGAGCCTGGTCCAAAGGGCACCCAAGGAGTTCCTGGAGCCCCTGGACCACCTGGA
Human_COL4A3            CAAGGAGAACCTGGTCTCCAGGGCACGCAAGGAGTTCCTGGGGCCCCCGGACCACCCGGA
Mouse_COL4A3            CAAGGAGAGCCAGGTCCAAAGGGAGCCCAAGGAGTCCCCGGCGTCCTGGGACCACCTGGA
                        ** ***** ** ****   ****  * ******** ** ** * **  ******** ***

Canine_Normal_5-18      GAAGCCGGTCCTAAGGGAGAATTCGGTATTTCGACACCAGTCCCAGGGCCCCCAGGACCT
Human_COL4A3            GAAGCCGGCCCTAGGGGAGAGCTCAGTGTTTCAACACCAGTTCCAGGCCCTCCAGGACCT
Mouse_COL4A3            GAAGCCGGTCTTAAAGGAGAACCCAGTACATCAACTCCAGATCTAGGTCCCCCAGGCCCT
                        ******** * **  *****   * **   ** ** ****  * *** ** ***** ***

Canine_Normal_5-18      CCAGGGCCCCCTGGTTATGCTGGCCCCCGAGGTCCACCTGGTATCCCTGGATCCATAGGA
Human_COL4A3            CCAGGGCCCCCTGGCCATCCTGGCCCCCAAGGTCCACCTGGTATCCCTGGATCCCTGGGG
Mouse_COL4A3            CCAGGGCCCCCTGGTCAGGCTGGCCCCAGAGGTCTACCTGGTTTGCCTGGACCGGTGGGA
                        **************  *  ********  ***** ******* * ****** *  * **

Canine_Normal_5-18      AAATGTGGTGATCCGGGTCTTCCTGGGCCTGATGGTGAACCAGGAATTCCAGGAATTGGC
Human_COL4A3            AAATGTGGAGATCCTGGTCTTCCAGGGCCTGATGGTGAACCAGGAATTCCAGGAATTGGA
Mouse_COL4A3            AAATGTG---ATCCTGGTCTTCCTGGACCTGATGGTGAGCCAGGAATTCCAGAAGCTGGA
                        *******   **** ******** ** *********** ************* *  ***

Canine_Normal_5-18      TTCCCTGGGCCCCCTGGACCTAAGGGAGACCAAGGTTTTCCAGGAGCAAAAGGAGCACCG
Human_COL4A3            TTTCCTGGGCCACCTGGACCTAAGGGAGACCAAGGTTTTCCAGGTACAAAAGGATCACTG
Mouse_COL4A3            TGCCCAGGGCCTCCAGGACCTAAGGGAAATCAAGGTTTCCCAGGGACAAAAGGATCACCT
                        *  ** ***** ** ************ * ******** *****  ******** ***

Canine_Normal_5-18      GGTTGTCCAGGAGAAATGGGGAAGCCCGGGTCACCTGGAGAACCAGGTCTCCCAGGAGCC
Human_COL4A3            GGTTGTCCTGGAAAAATGGGAGAGCCTGGGTTACCTGGAAAGCCAGGCCTCCCAGGAGCC
Mouse_COL4A3            GGTTGTCCTGGGGAAATGGGAAAGCCTGGCCGTCCTGGAGAACCAGGCATCCCAGGAGCC
                        ******** **  *******  **** **    ****** * *****  ***********
```

```
Canine_Normal_5-18      AAGGGAGAACCAGGACTAGCCATACCTGGAGAACCAGGAATACCAGGTTTTCCAGGAGAA
Human_COL4A3            AAGGGAGAACCAGCAGTAGCCATGCCTGGAGGACCAGGAACACCAGGTTTTCCAGGAGAA
Mouse_COL4A3            AAGGGAGAACCATCCGTAGGCAGGCCTGGGAAACCAGGGAAACCAGGCTTTCCAGGAGAA
                        ************    *** **  *****   ****** * ****** ************

Canine_Normal_5-18      AGAGGCAATTCTGGGGAAAATGGAGAAATTGGACTCCCTGGACTTCCAGGTCTCCCTGGA
Human_COL4A3            AGAGGCAATTCTGGGGAACATGGAGAAATTGGACTCCCTGGACTTCCAGGTCTCCCTGGA
Mouse_COL4A3            AGAGGCAATGCCGGGGAAAATGGAGATATTGGACTCCCTGGGCTTCCAGGCCTCCCTGGA
                        ********* * ****** ******* ************** ******** *********

Canine_Normal_5-18      ATTCCAGGAACTGGAGGGCTTGATGGACCACCAGGGGATCCAGGGAAGCCTGGACCACCT
Human_COL4A3            ACTCCAGGAAATGAAGGGCTTGATGGACCACGAGGAGATCCAGGGCAGCCTGGACCACCT
Mouse_COL4A3            ACTCCAGGAAGAGGCGGGCTTGACGGGCCTCCAGGAGACCCAGGACAGCCTGGATCACCT
                        * ********  *  ******** ** ** * *** ** *****  ******** *****

Canine_Normal_5-18      GGAGAAAAAGGACCCCCAGGAAGTTGCACAGAGGGTCCCAGAGGAGCCCACGGACTTCCA
Human_COL4A3            GGAGAACAAGGACCCCCAGGAAGGTGCATAGAGGGTCCCAGGGGAGCCCAAGGACTTCCA
Mouse_COL4A3            GGAGCAAAAGGATCCCCAGGAAGGTGCATACCAGGACCCAGGGGCACCCAAGGACTTCCC
                        **** * ***** ********** **** *   ** ***** **  **** ********

Canine_Normal_5-18      GGATTAAATGGATTGGAAGGGCAACCAGGCAGAAGAGGTGAAACAGGGCCAAAAGGAGAC
Human_COL4A3            GGCTTAAATGGATTGAAAGGGCAACAAGGCAGAAGAGGTAAAACGGGGCCAAAGGGAGAC
Mouse_COL4A3            GGCTTGAATGGATTGAAAGGGCAACCAGGTCGAAGAGGTGATACCGGGCCAAAGGGAGAC
                        ** ** ********* ********* ***  ******** * ** ******** ******

Canine_Normal_5-18      CCAGGTATTCCAGGCTTGGATAGATCAGGCTTTCCTGGAGAACCTGGACCACCAGGAATG
Human_COL4A3            CCAGGAATTCCAGGCTTGGATAGATCAGGATTTCCTGGAGAAACTGGATCACCAGGAATT
Mouse_COL4A3            CCTGGCATCCCAGGCATGGACAGGTCTGGGGTCCCCGGAGATCCTGGACCACCAGGAACA
                        ** ** ** ****** **** ** ** **  * ** *****  ***** *********

Canine_Normal_5-18      CCAGGTCATCGAGGTGAGATAGGACCACCTGGCCAAAAAGGATATCCAGGAAATCCGGGA
Human_COL4A3            CCAGGTCATCAAGGTGAAATGGGACCACTGGGTCAAAGAGGATATCCAGGAAATCCGGGA
Mouse_COL4A3            CCAGGTTGTCCTGGTGAGATGGGACCACCTGGACAAAAAGGATATCCGGGAGCTCCAGGA
                        ******  **  ***** ** *******  ** **** ********* ***  *** ***

Canine_Normal_5-18      TTTTTAGGACTACCAGGTGAAGATGGAGTGATTGGGATGATGGGCTATCCAGGAAACACT
Human_COL4A3            ATTTTAGGGCCACCAGGTGAAGATGGAGTGATTGGGATGATGGGATTTCCTGGAGCCATT
Mouse_COL4A3            TTCCCAGGACCACCAGGCGAGAAAGGAGAGGTTGGAATGATGGGCTATCCTGGAACCACT
                         *   *** * ****** **  * **** * **** ******** * *** ***  ** *

Canine_Normal_5-18      GGCCCTCCTGGGCCTCCCGGGATCCCAGGCACCCCAGGACAGAGGGGTAGCTTTGGAATT
Human_COL4A3            GGCCCTCCAGGGCCCCCTGGGAACCCAGGCACACCAGGGCAGAGGGGGAGCCCTGGAATT
Mouse_COL4A3            GGCCCTCCGGGGCTTCCTGGGAAACCAGGCTCACAGGGGCAGCGAGGTAGCCTCGGAATC
                        ******** ****  ** ****  ****** * *  ** *** * ** ***   *****

Canine_Normal_5-18      CCAGGTGCAAAGGGTGAGAAAGGGCCCCCAGGAGCCAAGGGGGAAGAAGGAGAGAAAGGA
Human_COL4A3            CCAGGAGTAAAGGGCCAGAGAGGAACCCCAGGAGCCAAGGGGGAACAAGGAGATAAAGGA
Mouse_COL4A3            CCAGGAATGAAAGGGGAAAAAGGACGCCCAGGAGCCAAAGGCGAACGAGGAGAGAAAGGA
                        *****    ** **  * * ***   ************ ** ***  ****** ******

Canine_Normal_5-18      CCTATGGGGCCTTGTCAAATAACCAGCTCAGTGGGGGACAAAGGAGAACCAGGCCTCAAA
Human_COL4A3            AATCCCGGGCCTTCAGAGATATCCCACGTAATAGGGGACAAAGGAGAACCAGGTCTCAAA
Mouse_COL4A3            AAACCAGGGCCTTCTCAAACAACACTCTTAAAGGGAGACAAAGGAGAGCCTGGACTTAAA
                              *******   * * * *   *  *   ** *********** ** ** ** ***

Canine_Normal_5-18      GGGTTTGCAGGAAAGCCGGGTGAAAAAGGAAACAGAGGCATTCCAGGGTTACCAGGTTTG
Human_COL4A3            GGATTCGCAGGAAATCCAGGTGAGAAAGGAAACAGAGGCGTTCCAGGGATGCCAGGTTTA
Mouse_COL4A3            GGATTTGTTGGGAATCCAGGTGAGAAAGGAAACAGAGGCAACCCAGGGTTACCAGGTCCG
                        ** ** *  ** ** ** ***** ***************   ****** * ******

Canine_Normal_5-18      AAAGGATTCGAAGGGCCACCTGGACCACCAGGCCCACCAGGCCCCCGAGGAGATCCAGGC
Human_COL4A3            AAGGGCCTCAAAGGACTACCCGGACCAGCAGGACCACCAGGCCCCAGAGGAGATTTGGGC
Mouse_COL4A3            AAAGGCCTCGAGGGATTGCCTGGGCTACCAGGTCCTCCAGGCCCTAGAGGAGATACGGGA
                        ** **  ** * **    ** ** * * **** ** ********  ********   **
```

```
Canine_Normal_5-18      AGCATTGGGAATCCTGGAGAAGCAGGACCACGTGGAAGTCCAGGAAACATGGGGAACATG
Human_COL4A3            AGCACTGGGAATCCTGGAGAACCAGGACTCCGTGGTATACCAGGAAGCATGGGGAACATG
Mouse_COL4A3            AGCAGAGGAAATCCTGGAAGACCAGGACCACACGGCATGCCAGGAAGCATGGGGATCATG
                        ****  ** *********  * ******  *  ** *  ******* ******** ****

Canine_Normal_5-18      GGGATGCCAGGTCTTAAAGGACACAGGGGACCTTTGGGACTACCAGGTGTCACTGGAAGA
Human_COL4A3            GGCATGCCAGGTTCTAAAGGAAAAAGGGGAACTTTGGGATTCCCAGGTCGAGCAGGAAGA
Mouse_COL4A3            GGGGTGCCAGGCCCTAAAGGGAGAAAAGGAACTTCAGGGCTTCCAGGTCTAGCTGGAAGA
                        **  *******   ******    *  *** ***  **  * ******    * ******

Canine_Normal_5-18      CCAGGCCTCCCAGGTGTTCACGGTCTCCAAGGAGACAAGGGAGAGCCAGGTTATTCAGCA
Human_COL4A3            CCAGGCCTCCCAGGTATTCATGGTCTCCAGGGAGATAAGGGAGAGCCAGGTTATTCAGAA
Mouse_COL4A3            CCAGGCCGGACAGGGATCCATGGTCCCCAAGGAGATAAGGGGGAGCCGGGTTATTCAGAA
                        *******   ****  * ** **** *** ***** ***** ***** ********** *

Canine_Normal_5-18      GGTACAAGGCCAGGACCACCGGGACTAAAGGGAGATCCAGGATTACCAGGTGCCATGGGA
Human_COL4A3            GGTACAAGGCCAGGACCACCGGGACCAACGGGGGATCCAGGACTGCCGGGTGATATGGGA
Mouse_COL4A3            GGTGCAAGACCAGGACCACCAGGACCAAAGGGCGATCCAGGATTGCCAGGTGACAAAGGA
                        *** **** *********** **** ** *** ********* * ** ****  *  ***

Canine_Normal_5-18      AGGAAAGGAGAAAGAGGGCTACCTGGCACCCCTGGCCATTCGGGGCCTGCTGGAACTGAG
Human_COL4A3            AAGAAAGGAGAAATGGGGCAACCTGGCCCACCTGGACATTTGGGGCCTGCTGGACCTGAG
Mouse_COL4A3            AAGAAAGGAGAAAGAGGCGTACCTGGGCCACCTGGACAATCGGGGCCTGCCGGACCTGAT
                        * ***********  **   ******  * ***** ** * ********* *** ****

Canine_Normal_5-18      GGAGCCCCTGGACATCCCGGAAGTCCTGGCCACCCAGGAAAGCCAGGCCCTGATGGAGAT
Human_COL4A3            GGAGCCCCTGGAAGTCCTGGAAGTCCTGGCCTCCCAGGAAAGCCAGGTCCTCATGGTGAT
Mouse_COL4A3            GGAGCCCCTGGGAGTCCTGGGAGTCCTGGTCACCCAGGAAAACCAGGTCCTGCTGGTGAT
                        ***********   *** ** ******** * ********* ***** ***  *** ***

Canine_Normal_5-18      TTGGGGTTAAAAGGCATCAAAGGCTTCCCTGGGTCCCCAGGAGTCAAAGGACCTCCAGGA
Human_COL4A3            TTGGGTTTTAAAGGAATCAAAGGCCTCCTGGGCCCTCCAGGAATCAGAGGCCCTCCAGGT
Mouse_COL4A3            TTGGGTCTGAAAGGACAGAAAGGCTTCCCAGGCCCTCCAGGAAGCACTGGCCCTCCAGGC
                        *****  * *****    ****** ***  **  * ******  **  ** ********

Canine_Normal_5-18      CCTCCAGGAATCCTAGGATCTCCTGGACCTATGGGCATGAGAGGCAACCAGGGCCGTGAT
Human_COL4A3            CTTCCAGGATTTCCAGGATCTCCTGGACCAATGGGTATAAGAGGTGACCAAGGACGTGAT
Mouse_COL4A3            CCTCCAGGACTCCCAGGACTTCCTGGGCCAATGGGTATGAGAGGTGACCAAGGACGTGAT
                        * ******* * * ****  ****** ** ***** ** *****  **** ** ******

Canine_Normal_5-18      GGAATCCCTGGTCCCCCAGGAGAAAAGGGAGAAACAGGTTTGCTGGGGGCACATCCAGGC
Human_COL4A3            GGAATTCCTGGTCCAGCCGGAGAAAAGGGAGAAACGGGTTTATTGAGGGCCCCTCCAGGC
Mouse_COL4A3            GGAATTCCTGGTCCTCCAGGAGAAAAAGGAGAAACAGGCTTGCTGGGGGCCTACCCAGGC
                        ***** ********  * ******** ******** ** **  ** ****    ******

Canine_Normal_5-18      CCAGGAGGGAGCCCTGGTCCTCCAGGAGCCAAAGGAGACAGGGGCGCCCCAGGCCTACCA
Human_COL4A3            CCAAGAGGGAACCCTGGTGCTCAAGGAGCCAAAGGAGACAGGGGAGCCCCAGGTTTTCCT
Mouse_COL4A3            CCAAAAGGGAGCCCTGGTGTACCAGGTGCCAAAGGAGACAGGGGAGTTCCAGGCTTGTCT
                        ***  ***** *******   * *** ***************** *  *****  *  *

Canine_Normal_5-18      GGCCCCCCAGGCAGGAAAGGGGCAGTGGGAGATGCCGGGCCACGGGGACCCACTGGCATG
Human_COL4A3            GGCCTCCCGGGCAGAAAAGGGGCCATGGGAGATGCTGGACCTCGAGGACCCACAGGCATA
Mouse_COL4A3            GGCCTTCCCGGCAGGAAGGGGGTTATGGGGGATGTTGGACCCCAAGGACCCCCAGGCACT
                        ****  ** ***** ** ****   **** ****  ** ** *  ****** * ****

Canine_Normal_5-18      ACGGGACTCCCAGGGCCACCAGGTTTTCCTGGTGCAGTCATCCCTGGCCAGAAAGGAAAT
Human_COL4A3            GAAGGATTCCCAGGGCCACCAGGTCTGCCCGGTGCAATTATCCCTGGCCAGACAGGAAAT
Mouse_COL4A3            GCTGGACTCCCAGGGCCACCAGGTCTACCTGGGGCAATTATCCCTGGCCCCAAAGGAGAC
                           *** ***************** * ** ** *** * **********  * **** *

Canine_Normal_5-18      CAAGGTCCACCAGGCTTCAGAGGAAACCCAGGTGAGCCTGGTCCTCCAGGCCCTCCAGGG
Human_COL4A3            CGTGGTCCACCAGGCTCAAGAGGAAGCCCAGGTGCGCCTGGTCCCCCTGGACCTCCAGGG
Mouse_COL4A3            AGAGGTCTTCCTGGCTTAAGAGGAAATCCAGGTGAGCCAGGTCCCCCTGGACCTCCAGGA
                           ****  ** ****  *******  ******* *** ***** ** ** ********
```

```
Canine_Normal_5-18      AGCCACGTGAAAGGCATAAAAGGAGACAAGGGACTCATGGGTGAGCCTGGCCCCAGAGGT
Human_COL4A3            AGTCATGTAATAGGCATAAAAGGAGACAAAGGGTCTATGGGCCACCCTGGCCCAAAAGGT
Mouse_COL4A3            CCTATCGGAAAAGGCATAAAAGGTGACAAAGGATTTATGGGCCCACCTGGCCCCAAAGGC
                              *  * ************ ***** **    *****    ******** * ***

Canine_Normal_5-18      CTACCTGGAACTGTGGGAGACAAGGGGCCACCGGGTCAACCGGGAGCACCAGGGAGCCCA
Human_COL4A3            CCACCTGGAACTGCAGGAGACATGGGACCACCAGGTCGTCTGGGAGCACCAGGTACTCCA
Mouse_COL4A3            CTGCCTGGAACTGTAGGGGACATGGGCCCACCAGGTTTCCCGGGAGCACCAGGTACCCCC
                        *  **********  ** **** *** ***** ***   * ************ *  **

Canine_Normal_5-18      GGTCTGCCAGGGCTCAGGGGCGATCCTGGATTCTATGGATTTCCAGGTGTTAAAGGAGAG
Human_COL4A3            GGTCTTCCAGGACCCAGAGGTGATCCTGGATTCCAGGGGTTTCCAGGCGTGAAAGGAGAA
Mouse_COL4A3            GGTCTTCCCGGTGTCAGAGGTGATCCAGGATTCCCTGGATTTCCAGGCATAAAAGGAGAA
                        ***** ** **   *** ** ***** ******   ** ********  * ********

Canine_Normal_5-18      AAGGGTAACCCAGGATTTCCGGGACCAGCTGGACCTCCAGGGCAAATTGGGCCAAAAGGA
Human_COL4A3            AAGGGTAATCCTGGATTTCTAGGATCCATTGGACCTCCAGGACCAATTGGGCCAAAAGGA
Mouse_COL4A3            AAGGGTAACCCTGGATTTCTTGGGCCAATTGGACATCCAGGACCAGTTGGGCCCAAAGGA
                        ******** ** *******  **  *   ***** ****** * * ******* ******

Canine_Normal_5-18      CCACCTGGTGTCCGTGGAGACCCTGGCACAGTTAAGATCATCTCCCTTCCAGGAAGCCCA
Human_COL4A3            CCACCTGGTGTACGTGGAGACCCTGGCACACTTAAGATTATCTCCCTTCCAGGAAGCCCA
Mouse_COL4A3            CCACCAGGTCCACGTGGAAAACCTGGCACTCTTAAGGTCATCTCTCTTCCGGGAAGCCCA
                        ***** ***   ****** * ********  ***** * ***** ***** *********

Canine_Normal_5-18      GGGCCACCTGGCCGTGCTGGAGGACCGGGGATGCCAGGAGAACCAGGGCCACCAGGGCCA
Human_COL4A3            GGGCCACCTGGCACACCTGGAGAACCAGGGATGCAGGGAGAACCTGGGCCACCAGGGCCA
Mouse_COL4A3            GGGCCACCTGGTGTACCTGGACAGCCAGGAATGAAAGGAGACCCTGGACCCCTGGGACTG
                        ***********     *****   ** ** ***   ***** ** ** ** *  ** *

Canine_Normal_5-18      CCAGGAATCCTAGGACCCTGTGGACCAAGAGGTAAACCAGGCATGGATGGAATACCAGGA
Human_COL4A3            CCTGGAAACCTAGGACCCTGTGGGCCAAGAGGTAAGCCAGGCAAGGATGGAAAACCAGGA
Mouse_COL4A3            CCAGGAATCCCAGGACCCTGTGGGCCAAGAGGTAAACCAGGCAAGGATGGGAAACCAGGA
                        ** **** ** ************ ************ ******* ****** * *******

Canine_Normal_5-18      ACTCCTGGGCCAATTGGAGAAAAAGGCAACAAAGGTTGTAAAGGCGAGCAAGGACCACCT
Human_COL4A3            ACTCCTGGACCAGCTGGAGAAAAAGGCAACAAAGGTTCTAAAGGAGAGCCAGGACCAGCT
Mouse_COL4A3            ACTCCAGGACCAGCTGGTACAAAGGGCAACAAGGGCTTGAAAGGACAGCAAGGCCCACCT
                        ***** ** ***  ***   *** ******** ** *  *****  *** *** *** **

Canine_Normal_5-18      GGATTAGATGGACTGCCAGGCTTGAAGGGGAGACCTGGAGACCCTGGACCACCTGCAACT
Human_COL4A3            GGATCAGATGGATTGCCAGGTTTGAAAGGAAAACGTGGAGACAGTGGATCACCTGCAACC
Mouse_COL4A3            GGTCTGGATGGATTGCCAGGCTTAAAGGGAAATCCTGGTGACAGGGGAACGCCGGCCACT
                        **    ****** ******* ** ** ** *  * *** ***   *** * ** ** **

Canine_Normal_5-18      GGGACCATGAAGAGGGGCTTTATCTTCACCAGGCATAGTCAGACCACAGTGATTCCTTCC
Human_COL4A3            TGGACAAC---GAGAGGCTTTGTCTTCACCCGACACAGTCAAACCACAGCAATTCCTTCA
Mouse_COL4A3            GGTACAAGAATGCGAGGCTTCATCTTCACCCGACACAGTCAAACCACGGCCATTCCTTCA
                         * ** *    * * *****  ******** * ** ***** ***** *  ********

Canine_Normal_5-18      TGTCCAGAAGGGACAGCGCCACTCTATAGTGGCTTTTCTCTTCTTTTCATACAAGGAAAT
Human_COL4A3            TGTCCAGAGGGGACAGTGCCACTCTACAGTGGGTTTTCTTTTCTTTTTGTACAAGGAAAT
Mouse_COL4A3            TGCCCTGAAGGAACACAGCCACTCTATAGTGGGTTTTCTCTTCTTTTTGTACAAGGAAAC
                        ** ** ** ** ***  ********** ***** ****** *******  **********

Canine_Normal_5-18      GAACAAGCCCATGGACAAGACTTGGGAACTCTCGGCAGCTGCCTGCAGCGATTTACCACA
Human_COL4A3            CAACGAGCCCACGGACAAGACCTTGGAACTCTTGGCAGCTGCCTGCAGCGATTTACCACA
Mouse_COL4A3            AAACGTGCACATGGACAAGACCTAGGTACTCTGGGCAGCTGCCTGCAGCGATTCACCACA
                         ***  ** ** ********* * ** ***** ******************** ******

Canine_Normal_5-18      ATGCCATTCTTGTTCTGTAACATCAACAATGTCTGTAATTTTGCATCTCGAAATGACTAC
Human_COL4A3            ATGCCATTCTTATTCTGCAATGTCAATGATGTATGTAATTTTGCATCTCGAAATGATTAT
Mouse_COL4A3            ATGCCGTTCTTATTCTGTAACATCAATAATGTATGTAACTTTGCATCACGAAATGATTAT
                        ***** ***** ***** **  ****  **** ***** ******** ******** **
```

```
Canine_Normal_5-18      TCATACTGGCTGTCAACACCAGCTCTGATGCCAATGGACATGGCTCCAATTACTGGCAGG
Human_COL4A3            TCATACTGGCTGTCAACACCAGCTCTGATGCCAATGAACATGGCTCCCATTACTGGCAGA
Mouse_COL4A3            TCATACTGGCTGTCAACACCAGCTCTGATGCCAATGGACATGGCTCCAATTAGTGGCAGA
                        ************************************ ********** **** ******

Canine_Normal_5-18      GCCCTGGAGCCTTATATTAGCAGATGCACTGTCTGTGAAGGTCCTACGATTGCCATAGCC
Human_COL4A3            GCCCTTGAGCCTTATATAAGCAGATGCACTGTTTGTGAAGGTCCTGCGATCGCCATAGCC
Mouse_COL4A3            GCTCTCGAACCCTATATTAGCAGATGCACCGTCTGTGAAGGTCCAGCAATGGCCATAGCT
                        ** ** ** ** ***** ********** ** **********  * ** ********

Canine_Normal_5-18      ATTCACAGCCAAACCACTGACATTCCCTCATGTCCCAATGGCTGGATTTCTCTCTGGAAA
Human_COL4A3            GTTCACAGCCAAACCACTGACATTCCTCCATGTCCTCACGGCTGGATTTCTCTCTGGAAA
Mouse_COL4A3            GTTCACAGTCAAACTACTGCTATCCCTCCGTGTCCCCAGGACTGGGTTTCTCTCTGGAAA
                         ******* ***** ****  ** **  * *****  * * **** **************

Canine_Normal_5-18      GGATTTTCGTTTATCATGTTCACAAGTGCAGGTTCTGAGGGTGCTGGGCAAGCACTGGCC
Human_COL4A3            GGATTTTCATTCATCATGTTCACAAGTGCAGGTTCTGAGGGCACCGGGCAAGCACTGGCC
Mouse_COL4A3            GGTTTTTCTTTCATTATGTTCACAAGTGCAGGCTCTGAGGGTGCTGGACAAGCACTTGCC
                        ** ***** ** ** ***************** ********  * ** ******** ***

Canine_Normal_5-18      TCCCCCGGCTCCTGCCTGGAAGAGTTCCGAGCCAGTCCGTTTATAGAATGTCACGGAAGA
Human_COL4A3            TCCCCTGGCTCCTGCCTGGAAGAATTCCGAGCCAGCCCATTTCTAGAATGTCATGGAAGA
Mouse_COL4A3            TCGCCTGGCTCCTGCCTGGAAGAATTCCGAGCCAGTCCATTTATAGAATGCCATGGACGA
                        ** ** ***************** ********** ** *** ******* ** *** **

Canine_Normal_5-18      GGAACGTGCAACTACTATTCAAATTCCTACAGTTTCTGGTTGGCTTCGTTAAACCCCCAA
Human_COL4A3            GGAACGTGCAACTACTATTCAAATTCCTACAGTTTCTGGCTGGCTTCATTAAACCCAGAA
Mouse_COL4A3            GGGACATGTAACTACTACTCAAACTCCTACAGTTTCTGGCTGGCTTCGCTGAACCCAGAA
                        ** ** ** ******** ***** *************** *******  * *****  **

Canine_Normal_5-18      AGAATGTTCAGAAAACCTATTCCATCAACTGTGAAAGCTG---AATCCATATGACTAGT-
Human_COL4A3            AGAATGTTCAGAAAGCCTATTCCATCAACTGTGAAAGCTGGGGAATTAGAAAAAATAATA
Mouse_COL4A3            AGAATGTTCAGAAAACCTATTCCATCAACTGTGAAAGCTGGAGACTTAGAGAAAATCATA
                        ************** *************************   * *       * *  *

Canine_Normal_5-18      -----------------------------------
Human_COL4A3            AGTCGCTGTCAGGTGTGCATGAAGAAAAGACACTGA
Mouse_COL4A3            AGCCGCTGTCAGGTGTGCATGAAGAAAAGACATTGA
```

APPENDIX 3

Canine COL4A3 amino acid sequence alignments

```
Consensus key (see documentation for details)
* - single, fully conserved residue
: - conservation of strong groups
. - conservation of weak groups
  - no consensus


CLUSTAL W (1.81) multiple sequence alignment


Canine_A3_Normal_5-18_Transl      --------------------------------------------------
Human_COL4A3_Translated_-_Fr      MSARTAPRPQVLLLPLLLVLLAAAPAASKGCVCKDKGQCFCDGAKGEKGE
Mouse_COL4A3_Translated_-_Fr      MHSKTAPRFLVFLLLTLLLLLAASPVASKGCVCKGKGQCLCAGTKGEKGE


Canine_A3_Normal_5-18_Transl      --------------------------------------------------
Human_COL4A3_Translated_-_Fr      KGFPGPPGSPGQKGFTGPEGLPGPQGPKGFPGLPGLTGSKGVRGISGLPG
Mouse_COL4A3_Translated_-_Fr      KGVPGSPGFPGQKGFPGPEGLPGPQGPKGSPGLPGLTGPKGIRGITGLPG


Canine_A3_Normal_5-18_Transl      ----------NFKKEIYHGDGVPGCNGSKGEQGFPGLPGTPGYPGIPGAV
Human_COL4A3_Translated_-_Fr      FSGSPGLPGTPGNTGPYGLVGVPGCSGSKGEQGFPGLPGTPGYPGIPGAA
Mouse_COL4A3_Translated_-_Fr      FAGPPGLPGLPGHPGPRGLAGLPGCNGSKGEQGFPGFPGTPGYAGLPGPD
                                             :       *:***.**********:******.*:**.

Canine_A3_Normal_5-18_Transl      GLKGEKGAPATGEGIELDGRGDPGLPGAPGFQGLPGLPGFPGPAGPPGPP
Human_COL4A3_Translated_-_Fr      GLKGQKGAPAKGEDIELDAKGDPGLPGAPGPQGLPGPPGFPGPVGPPGPP
Mouse_COL4A3_Translated_-_Fr      GLKGQKGEPAQGEDRGFNGKGDPGPPGVPGFQGFPGLPGFPGPAGPPGPP
                                  ****:** **  *.  ::.:**** **.** **:** ******.******

Canine_A3_Normal_5-18_Transl      GFLGFPGAMGPPGPKGQMGDKVIGQKGERGVKGLTGPPGPPGTVIVTLTG
Human_COL4A3_Translated_-_Fr      GFFGFPGAMGPRGPKGHMGERVIGHKGERGVKGLTGPPGPPGTVIVTLTG
Mouse_COL4A3_Translated_-_Fr      GFFGLPGAMGPRGPKGHMGDSVIGQKGERGMKGLTGPPGPPGTVIFTLTQ
                                  **:*:****** ****:**: ***:*****:**************.***

Canine_A3_Normal_5-18_Transl      PDNRTDLKGEKGDKGAMGKPGPPGPLGPPGESYGSEKGAPGEPGPQGKPG
Human_COL4A3_Translated_-_Fr      PDNRTDLKGEKGDKGAMGEPGPPGPSGLPGESYGSEKGAPGDPGLQGKPG
Mouse_COL4A3_Translated_-_Fr      PYNKSDFKGEKGDEGERGEPGPPGPSGPPGDSYGSEKGAPGEPGPRGKPG
                                  * *::*:******:*  *:****** * **:**********:** :****

Canine_A3_Normal_5-18_Transl      KDGAPGFPGTEGPKGNRGFPGLRGEDGIKGWKGDIGPPGFRGPTEYYDAY
Human_COL4A3_Translated_-_Fr      KDGVPGFPGSEGVKGNRGFPGLMGEDGIKGQKGDIGPPGFRGPTEYYDTY
Mouse_COL4A3_Translated_-_Fr      KDGAPGFPGTEGAKGNRGFPGLRGEAGIKGRKGDIGPPGFPGPTEYYDAY
                                  ***.*****:** ********* ** **** ********* *******:*

Canine_A3_Normal_5-18_Transl      QEKGDEGIPGPPGPKGARGPQGPSGPPGVLGSAGSSKPGLRGAPGSPGMK
Human_COL4A3_Translated_-_Fr      QEKGDEGTPGPPGPRGARGPQGPSGPPGVPGSPGSSRPGLRGAPGWPGLK
Mouse_COL4A3_Translated_-_Fr      LEKGERGMPGLPGPKGARGPQGPSGPPGVPGSPGLSRPGLRGPIGWPGLK
                                   ***:.* ** ***:************** **.* *:*****. * **:*

Canine_A3_Normal_5-18_Transl      GRKGEQGPPGKNAVGPPGSPGCPGSPGPIGLPGYPGPPGDIVFRKGPPGV
Human_COL4A3_Translated_-_Fr      GSKGERGRPGKDMGTPGSPGCAGSPGLPGSPGPPGPPGDIVFRKGPPGD
Mouse_COL4A3_Translated_-_Fr      GSKGERGPPGKDTVGPPGPLGCPGSPGPPGPPGPPGCPGDIVFKCSP-GE
                                  * ***:* ***:::*.**. **.****  * ** ** ******: .* *

Canine_A3_Normal_5-18_Transl      GGLPGHVGFPGIPGVEGPKGEPGLLCVQCAYIPGRPGLPGLPGLDGIKGI
Human_COL4A3_Translated_-_Fr      HGLPGYLGSPGIPGVDGPKGEPGLLCTQCPYIPGPPGLPGLPGLHGVKGI
Mouse_COL4A3_Translated_-_Fr      HGMPGDTGPPGVPGLDGPKGEPGSPCTECHCFPGPPGVPGFPGLDGIKGI
                                   *:**  * **:**::*******  *.:*  :** **:**:***.*:***
```

```
Canine_A3_Normal_5-18_Transl      PGGQGAAGIKGSPGSPGSAGLPGFPGFPGAPGAPGLKGEKGETSEPEGQV
Human_COL4A3_Translated_-_Fr      PGRQGAAGLKGSPGSPGNTGLPGFPGFPGAQGDPGLKGEKGETLQPEGQV
Mouse_COL4A3_Translated_-_Fr      PGGRGVPGLKGNPGSPGSAGLPGFAGFPGDQGHPGLKGDKGDTPLPWGQV
                                  ** :*..*:**.*****.:*****.****  * *****:**:*  * ***

Canine_A3_Normal_5-18_Transl      GAPGDPGLRGHPGRKGLDGIPGTPGIKGPPGPKGEPALSGEKGDQGPPGD
Human_COL4A3_Translated_-_Fr      GVPGDPGLRGQPGRKGLDGIPGTLGVKGLPGPKGELALSGEKGDQGPPGD
Mouse_COL4A3_Translated_-_Fr      GNPGDPGLRGLPGRKGFDGTPGGPGAKGPPGPQGEPALSGRKGDQGPPGP
                                  * ******** *****:** **  * ** ***:** ****.********

Canine_A3_Normal_5-18_Transl      PGIPGSPGPAGPPGPPNYGPQGEPGPKGTQGVPGAPGPPGEAGPKGEFGI
Human_COL4A3_Translated_-_Fr      PGSPGSPGPAGPAGPPGYGPQGEPGLQGTQGVPGAPGPPGEAGPRGELSV
Mouse_COL4A3_Translated_-_Fr      PGFPGPPGPAGPAGPPGYGPQGEPGPKGAQGVPGVLGPPGEAGLKGEPST
                                  ** **.******.***.******** :*:*****. ******* :** .

Canine_A3_Normal_5-18_Transl      STPVPGPPGPPGPPGYAGPRGPPGIPGSIGKCGDPGLPGPDGEPGIPGIG
Human_COL4A3_Translated_-_Fr      STPVPGPPGPPGPPGHPGPQGPPGIPGSLGKCGDPGLPGPDGEPGIPGIG
Mouse_COL4A3_Translated_-_Fr      STPDLGPPGPPGPPGQAGPRGLPGLPGPVGKC-DPGLPGPDGEPGIPEAG
                                  ***  ********** .**:* **:**.:*** **************  *

Canine_A3_Normal_5-18_Transl      FPGPPGPKGDQGFPGAKGAPGCPGEMGKPGSPGEPGLPGAKGEPGLAIPG
Human_COL4A3_Translated_-_Fr      FPGPPGPKGDQGFPGTKGSLGCPGKMGEPGLPGKPGLPGAKGEPAVAMPG
Mouse_COL4A3_Translated_-_Fr      CPGPPGPKGNQGFPGTKGSPGCPGEMGKPGRPGEPGIPGAKGEPSVGRPG
                                   ********:*****:**: ****:**:** **:**:*******.:. **

Canine_A3_Normal_5-18_Transl      EPGIPGFPGERGNSGENGEIGLPGLPGLPGIPGTGGLDGPPGDPGKPGPP
Human_COL4A3_Translated_-_Fr      GPGTPGFPGERGNSGEHGEIGLPGLPGLPGTPGNEGLDGPRGDPGQPGPP
Mouse_COL4A3_Translated_-_Fr      KPGKPGFPGERGNAGENGDIGLPGLPGLPGTPGRGGLDGPPGDPGQPGSP
                                   ** *********:**:*:********** **  ***** ****:**.*

Canine_A3_Normal_5-18_Transl      GEKGPPGSCTEGPRGAHGLPGLNGLEGQPGRRGETGPKGDPGIPGLDRSG
Human_COL4A3_Translated_-_Fr      GEQGPPGRCIEGPRGAQGLPGLNGLKGQQGRRGKTGPKGDPGIPGLDRSG
Mouse_COL4A3_Translated_-_Fr      GAKGSPGRCIPGPRGTQGLPGLNGLKGQPGRRGDTGPKGDPGIPGMDRSG
                                  * :*.** *  ****::********:** ****.***********:****

Canine_A3_Normal_5-18_Transl      FPGEPGPPGMPGHRGEIGPPGQKGYPGNPGFLGLPGEDGVIGMMGYPGNT
Human_COL4A3_Translated_-_Fr      FPGETGSPGIPGHQGEMGPLGQRGYPGNPGILGPPGEDGVIGMMGFPGAI
Mouse_COL4A3_Translated_-_Fr      VPGDPGPPGTPGCPGEMGPPGQKGYPGAPGFPGPPGEKGEVGMMGYPGTT
                                  .**:.*.** **  **:** **:**** **: * ***.* :****:**

Canine_A3_Normal_5-18_Transl      GPPGPPGIPGTPGQRGSFGIPGAKGEKGPPGAKGEEGEKGPMGPCQITSS
Human_COL4A3_Translated_-_Fr      GPPGPPGNPGTPGQRGSPGIPGVKGQRGTPGAKGEQGDKGNPGPSEISHV
Mouse_COL4A3_Translated_-_Fr      GPPGLPGKPGSQGQRGSLGIPGMKGEKGRPGAKGERGEKGKPGPSQTTLL
                                  **** ** **: ***** **** **::* ******.*:**  **.: :

Canine_A3_Normal_5-18_Transl      VGDKGEPGLKGFAGKPGEKGNRGIPGLPGLKGFEGPPGPPGPPGPRGDPG
Human_COL4A3_Translated_-_Fr      IGDKGEPGLKGFAGNPGEKGNRGVPGMPGLKGLKGLPGPAGPPGPRGDLG
Mouse_COL4A3_Translated_-_Fr      KGDKGEPGLKGFVGNPGEKGNRGNPGLPGPKGLEGLPGLPGPPGPRGDTG
                                   ***********.*:******** **:** **::* ** .******** *

Canine_A3_Normal_5-18_Transl      SIGNPGEAGPRGSPGNMGNMGMPGLKGHRGPLGLPGVTGRPGLPGVHGLQ
Human_COL4A3_Translated_-_Fr      STGNPGEPGLRGIPGSMGNMGMPGSKGKRGTLGFPGRAGRPGLPGIHGLQ
Mouse_COL4A3_Translated_-_Fr      SRGNPGRPGPHGMPGSMGIMGVPGPKGRKGTSGLPGLAGRPGRTGIHGPQ
                                  * ****..* :* **.** **:** **::*. *:** :**** .*:** *

Canine_A3_Normal_5-18_Transl      GDKGEPGYSAGTRPGPPGLKGDPGLPGAMGRKGERGLPGTPGHSGPAGTE
Human_COL4A3_Translated_-_Fr      GDKGEPGYSEGTRPGPPGPTGDPGLPGDMGKKGEMGQPGPPGHLGPAGPE
Mouse_COL4A3_Translated_-_Fr      GDKGEPGYSEGARPGPPGPKGDPGLPGDKGKKGERGVPGPPGQSGPAGPD
                                  ********* *:****** .*******  *:*** * **.**: ****.:

Canine_A3_Normal_5-18_Transl      GAPGHPGSPGHPGKPGPDGDLGLKGIKGFPGSPGVKGPPGPPGILGSPGP
Human_COL4A3_Translated_-_Fr      GAPGSPGSPGLPGKPGPHGDLGFKGIKGLLGPPGIRGPPGLPGFPGSPGP
Mouse_COL4A3_Translated_-_Fr      GAPGSPGSPGHPGKPGPAGDLGLKGQKGFPGPPGSTGPPGPPGLPGLPGP
                                  **** ***** ****** ****:** **: *.**  **** **: * ***
```

```
Canine_A3_Normal_5-18_Transl      MGMRGNQGRDGIPGPPGEKGETGLLGAHPGPGGSPGPPGAKGDRGAPGLP
Human_COL4A3_Translated_-_Fr      MGIRGDQGRDGIPGPAGEKGETGLLRAPPGPRGNPGAQGAKGDRGAPGFP
Mouse_COL4A3_Translated_-_Fr      MGMRGDQGRDGIPGPPGEKGETGLLGAYPGPKGSPGVPGAKGDRGVPGLS
                                  **:**:*********.********* * *** *.**  *******.**:.

Canine_A3_Normal_5-18_Transl      GPPGRKGAVGDAGPRGPTGMTGLPGPPGFPGAVIPGQKGNQGPPGFRGNP
Human_COL4A3_Translated_-_Fr      GLPGRKGAMGDAGPRGPTGIEGFPGPPGLPGAIIPGQTGNRGPPGSRGSP
Mouse_COL4A3_Translated_-_Fr      GLPGRKGVMGDVGPQGPPGTAGLPGPPGLPGAIIPGPKGDRGLPGLRGNP
                                  * *****.:**.**:**.*  *:*****:***:*** .*::* ** **.*

Canine_A3_Normal_5-18_Transl      GEPGPPGPPGSHVKGIKGDKGLMGEPGPRGLPGTVGDKGPPGQPGAPGSP
Human_COL4A3_Translated_-_Fr      GAPGPPGPPGSHVIGIKGDKGSMGHPGPKGPPGTAGDMGPPGRLGAPGTP
Mouse_COL4A3_Translated_-_Fr      GEPGPPGPPGPIGKGIKGDKGFMGPPGPKGLPGTVGDMGPPGFPGAPGTP
                                  * ********.   ******* ** ***:* ***.** ****  ****:*

Canine_A3_Normal_5-18_Transl      GLPGLRGDPGFYGFPGVKGEKGNPGFPGPAGPPGQIGPKGPPGVRGDPGT
Human_COL4A3_Translated_-_Fr      GLPGPRGDPGFQGFPGVKGEKGNPGFLGSIGPPGPIGPKGPPGVRGDPGT
Mouse_COL4A3_Translated_-_Fr      GLPGVRGDPGFPGFPGIKGEKGNPGFLGPIGHPGPVGPKGPPGPRGKPGT
                                  **** ****** ****:********* *. * ** :******* **.***

Canine_A3_Normal_5-18_Transl      VKIISLPGSPGPPGRAGGPGMPGEPGPPGPPGILGPCGPRGKPGMDGIPG
Human_COL4A3_Translated_-_Fr      LKIISLPGSPGPPGTPGEPGMQGEPGPPGPPGNLGPCGPRGKPGKDGKPG
Mouse_COL4A3_Translated_-_Fr      LKVISLPGSPGPPGVPGQPGMKGDPGPLGLPGIPGPCGPRGKPGKDGKPG
                                  :*:*********** .* *** *:*** * **  ********** ** **

Canine_A3_Normal_5-18_Transl      TPGPIGEKGNKGCKGEQGPPGLDGLPGLKGRPGDPGPPATGTMKRGFIFT
Human_COL4A3_Translated_-_Fr      TPGPAGEKGNKGSKGEPGPAGSDGLPGLKGKRGDSGSPATWTTR-GFVFT
Mouse_COL4A3_Translated_-_Fr      TPGPAGTKGNKGLKGQQGPPGLDGLPGLKGNPGDRGTPATGTRMRGFIFT
                                  **** * ***** **: **.* ********. ** *.*** *   **:**

Canine_A3_Normal_5-18_Transl      RHSQTTVIPSCPEGTAPLYSGFSLLFIQGNEQAHGQDLGTLGSCLQRFTT
Human_COL4A3_Translated_-_Fr      RHSQTTAIPSCPEGTVPLYSGFSFLFVQGNQRAHGQDLGTLGSCLQRFTT
Mouse_COL4A3_Translated_-_Fr      RHSQTTAIPSCPEGTQPLYSGFSLLFVQGNKRAHGQDLGTLGSCLQRFTT
                                  ******.******** *******:**:***::******************

Canine_A3_Normal_5-18_Transl      MPFLFCNINNVCNFASRNDYSYWLSTPALMPMDMAPITGRALEPYISRCT
Human_COL4A3_Translated_-_Fr      MPFLFCNVNDVCNFASRNDYSYWLSTPALMPMNMAPITGRALEPYISRCT
Mouse_COL4A3_Translated_-_Fr      MPFLFCNINNVCNFASRNDYSYWLSTPALMPMDMAPISGRALEPYISRCT
                                  *******:*:**********************:****:************

Canine_A3_Normal_5-18_Transl      VCEGPTIAIAIHSQTTDIPSCPNGWISLWKGFSFIMFTSAGSEGAGQALA
Human_COL4A3_Translated_-_Fr      VCEGPAIAIAVHSQTTDIPPCPHGWISLWKGFSFIMFTSAGSEGTGQALA
Mouse_COL4A3_Translated_-_Fr      VCEGPAMAIAVHSQTTAIPPCPQDWVSLWKGFSFIMFTSAGSEGAGQALA
                                  *****::***:***** **.**:.*:******************:*****

Canine_A3_Normal_5-18_Transl      SPGSCLEEFRASPFIECHGRGTCNYYSNSYSFWLASLNPQRMFRKPIPST
Human_COL4A3_Translated_-_Fr      SPGSCLEEFRASPFLECHGRGTCNYYSNSYSFWLASLNPERMFRKPIPST
Mouse_COL4A3_Translated_-_Fr      SPGSCLEEFRASPFIECHGRGTCNYYSNSYSFWLASLNPERMFRKPIPST
                                  **************:************************:**********

Canine_A3_Normal_5-18_Transl      VKA---ESIL-----------
Human_COL4A3_Translated_-_Fr      VKAGELEKIISRCQVCMKKRH
Mouse_COL4A3_Translated_-_Fr      VKAGDLEKIISRCQVCMKKRH
                                  ***   *.*:
```

APPENDIX 4

Canine COL4A4 nucleotide sequence

Sequence Range: 1 to 5163

```
                10          20          30          40          50
          CTAGA GGATC TACTA GTCAT ATGGA TGTGA GAGTC TTCTG GAAGA GAAGA
          GATCT CCTAG ATGAT CAGTA TACCT ACACT CTCAG AAGAC CTTCT CTTCT

                60          70          80          90         100
          GAAAC CTGAC AATCT TTTTT AAGTC CCAAG ACTTG ACTGT GAGGT AATCC
          CTTTG GACTG TTAGA AAAAA TTCAG GGTTC TGAAC TGACA CTCCA TTAGG

               110         120         130         140         150
          ATGCA CATGG CATCA GTAAA GTGTT CATTC AGGTG GATCA AGCCC TTGGC
          TACGT GTACC GTAGT CATTT CACAA GTAAG TCCAC CTAGT TCGGG AACCG

               160         170         180         190         200
          CAGAG ATCCC TGGTC ACTTA TATTT ATCCT CTTTT CTGTA CAACA TGCAT
          GTCTC TAGGG ACCAG TGAAT ATAAA TAGGA GAAAA GACAT GTTGT ACGTA

               210         220         230         240         250
          ATGGG AGTGG AAAGA AGTTT GTCGG CCCCT GTGGA GGAAG AGATT GCTCG
          TACCC TCACC TTTCT TCAAA CAGCC GGGGA CACCT CCTTC TCTAA CGAGC

               260         270         280         290         300
          GTGTG CCATT GCTTT CCTGA AAAAG GGTCT CGGGG TCAAC CAGGA CCGCC
          CACAC GGTAA CGAAA GGACT TTTTC CCAGA GCCCC AGTTG GTCCT GGCGG

               310         320         330         340         350
          GGGGC CACAG GGTCC CATTG GACCC CTGGG ACTGC CAGGA CCCAC AGGAA
          CCCCG GTGTC CCAGG GTAAC CTGGG GACCC TGACG GTCCT GGGTG TCCTT

               360         370         380         390         400
          TTCCA GGAGA GAAAG GGATG AGAGG TGACA GTGGC CCTCC TGGAG CAGCA
          AAGGT CCTCT CTTTC CCTAC TCTCC ACTGT CACCG GGAGG ACCTC GTCGT

               410         420         430         440         450
          GGTGA CAAAG GTGAC AAGGG TCCAA CTGGT GTTCC TGGAT TTCCA GGGTT
          CCACT GTTTC CACTG TTCCC AGGTT GACCA CAAGG ACCTA AAGGT CCCAA

               460         470         480         490         500
          GGATG GCATA CCTGG CCACC CGGGG CCTCC TGGAT CCAGA GGCAA GCCTG
          CCTAC CGTAT GGACC GGTGG GCCCC GGAGG ACCTA GGTCT CCGTT CGGAC

               510         520         530         540         550
          GCATG CATGG CTACA ATGGT TCACG AGGTG ATCCA GGGTT TCCAG GAAAA
          CGTAC GTACC GATGT TACCA AGTGC TCCAC TAGGT CCCAA AGGTC CTTTT

               560         570         580         590         600
          AGAGG AGTTC CTGGC CCAGG AGGCC CCCCA GGCCT TCCTG GGGAA AGGGG
          TCTCC TCAAG GACCG GGTCC TCCGG GGGGT CCGGA AGGAC CCCTT TCCCC

               610         620         630         640         650
          AGAAA AAGGA AATTC AGTGT TCATT TTAGG TGCCA TTAAA GGTAT TCAGG
          TCTTT TTCCT TTAAG TCACA AGTAA AATCC ACGGT AATTT CCATA AGTCC
```

```
        660         670         680         690         700
GAGAC AGGGG GGACC CAGAC CTGCC TGGCT TGCCG GGACC GAAGG GGGCA
CTCTG TCCCC CCTGG GTCTG GACGG ACCGA ACGGC CCTGG CTTCC CCCGT

        710         720         730         740         750
GGAGG ACCGG CAGGT CCCAA GGGAC AGCCA GGAGA GCCTG GGTTA ACAGG
CCTCC TGGCC GTCCA GGGTT CCCTG TCGGT CCTCT CGGAC CCAAT TGTCC

        760         770         780         790         800
TGCTC CGGGC CATCC TGGGA GACCA GGCTT GAAGG GTAAT CCTGG TGTGG
ACGAG GCCCG GTAGG ACCCT CTGGT CCGAA CTTCC CATTA GGACC ACACC

        810         820         830         840         850
GAGTA AAGGG GCAAA TGGGA GACCC GGGTG AGGTT GGCCA GCAGG GTTCT
CTCAT TTCCC CGTTT ACCCT CTGGG CCCAC TCCAA CCGGT CGTCC CAAGA

        860         870         880         890         900
CCTGG ACCCA CCTTA TTGGT ACAGC CACCT GATTC CTGTC TGTAT AAAGG
GGACC TGGGT GGAAT AACCA TGTCG GTGGA CTAAG GACAG ACATA TTTCC

        910         920         930         940         950
AGAAA AGGGC ATAAA AGGAA TGCCT GGTAT GATTG GACCT CCAGG ACTAC
TCTTT TCCCG TATTT TCCTT ACGGA CCATA CTAAC CTGGA GGTCC TGATG

        960         970         980         990        1000
CGGGA CCCAA GGGAG AACCT GGAAT TGGAC CAAAA GGAGA GAAGG GTATT
GCCCT GGGTT CCCTC TTGGA CCTTA ACCTG GTTTT CCTCT CTTCC CATAA

       1010        1020        1030        1040        1050
CCTGG GTTCT CAGGA CCTCG GGGTG ATCCG GGTTC CTATG GCTCT CCAGG
GGACC CAAGA GTCCT GGAGC CCCAC TAGGC CCAAG GATAC CGAGA GGTCC

       1060        1070        1080        1090        1100
TTTTC CAGGA TTAAA GGGGA AACCA GGACT GTTTG GAGAT CCTGG ATCAT
AAAAG GTCCT AATTT CCCCT TTGGT CCTGA CAAAC CTCTA GGACC TAGTA

       1110        1120        1130        1140        1150
TTGGA TTTGT TGGCC CAAAG GGGGA TCCTG GAGAC CGTGG GTACC CAGGA
AACCT AAACA ACCGG GTTTC CCCCT AGGAC CTCTG GCACC CATGG GTCCT

       1160        1170        1180        1190        1200
CCACC GGGTG TTCTG GTAAC TCCAT CTGTT CCACT CAAAG GCCCT CCAGG
GGTGG CCCAC AAGAC CATTG AGGTA GACAA GGTGA GTTTC CGGGA GGTCC

       1210        1220        1230        1240        1250
GGATC CGGGG CGCCC TGGCC ACTAT GGAGA AATGG GGTCC GTTGG ACCAC
CCTAG GCCCC GCGGG ACCGG TGATA CCTCT TTACC CCAGG CAACC TGGTG

       1260        1270        1280        1290        1300
CTGGT CCCCC TGGTC CCCCA GGTCC ACCAG GGGAA GACTG TGCAG GCATG
GACCA GGGGG ACCAG GGGGT CCAGG TGGTC CCCTT CTGAC ACGTC CGTAC

       1310        1320        1330        1340        1350
ATGGG ACCCC CTGGA CCAAG AGGAT TTCCT GGTCA TCCAG GATTT CCAGG
TACCC TGGGG GACCT GGTTC TCCTA AAGGA CCAGT AGGTC CTAAA GGTCC
```

```
       1360        1370        1380        1390        1400
GGCAG CTGGT ATCCC AGGGA GAGCG GATTT AAGTC CAGGA AAACC AGGGA
CCGTC GACCA TAGGG TCCCT CTCGC CTAAA TTCAG GTCCT TTTGG TCCCT

       1410        1420        1430        1440        1450
AACCG GGACC TCCTG GGTTG CCCGG AGCAC CAGGG CTGCA AGGAC CCCCA
TTGGC CCTGG AGGAC CCAAC GGGCC TCGTG GTCCC GACGT TCCTG GGGGT

       1460        1470        1480        1490        1500
GGATC AGATG TTATC TATTG TAGTG TTGGG CATCC TGGAC CACAA GGGAT
CCTAG TCTAC AATAG ATAAC ATCAC AACCC GTAGG ACCTG GTGTT CCCTA

       1510        1520        1530        1540        1550
AAAAG GCAAA ATGGG TCCTC CAGGA AGAAG AGGCT CGAAA GGAGA AAAAG
TTTTC CGTTT TACCC AGGAG GTCCT TCTTC TCCGA GCTTT CCTCT TTTTC

       1560        1570        1580        1590        1600
GAAAT GCGGG GCTCT GTGCC TGTGA GCCTG GTCCC ATGGG CCCAC CAGGC
CTTTA CGCCC CGAGA CACGG ACACT CGGAC CAGGG TACCC GGGTG GTCCG

       1610        1620        1630        1640        1650
CCTCC AGGAC TTCCT GGAAG GCAGG GTAGT AAGGG AGACT TGGGG CTCCC
GGAGG TCCTG AAGGA CCTTC CGTCC CATCA TTCCC TCTGA ACCCC GAGGG

       1660        1670        1680        1690        1700
TGGGT GGATT GGAGA GAAAG GTCAC CCAGG CCCTC CTGGT GCTGA AGGAT
ACCCA CCTAA CCTCT CTTTC CAGTG GGTCC GGGAG GACCA CGACT TCCTA

       1710        1720        1730        1740        1750
CTCCA GGATC ACCAG GAAAA CATGG CGCCT CAGGA CCACC TGGCA GCAAA
GAGGT CCTAG TGGTC CTTTT GTACC GCGGA GTCCT GGTGG ACCGT CGTTT

       1760        1770        1780        1790        1800
GGAGA AAAGG GCGAC AGGGT TGTAT CAAGA GTGAA AGGGA AAAAA GGAGA
CCTCT TTTCC CGCTG TCCCA ACATA GTTCT CACTT TCCCT TTTTT CCTCT

       1810        1820        1830        1840        1850
AAGAG GTCCT GATGG GCTCC CAGGA TTTCC AGGAC AACAG GGACA AAATG
TTCTC CAGGA CTACC CGAGG GTCCT AAAGG TCCTG TTGTC CCTGT TTTAC

       1860        1870        1880        1890        1900
GTCGG GATGG ACTTC CTGGA AAAAA GGGGG ATCCA GGCCC CCCGG GGGGT
CAGCC CTACC TGAAG GACCT TTTTT CCCCC TAGGT CCGGG GGGCC CCCCA

       1910        1920        1930        1940        1950
CATGA GGATG CAGCC CCAGG TGCTA AAGGG TCTCC TGGAC CACCG GGGCC
GTACT CCTAC GTCGG GGTCC ACGAT TTCCC AGAGG ACCTG GTGGC CCCGG

       1960        1970        1980        1990        2000
CCCAG GCAGA ACAGG ACCCA GGGGA CAACC AGGTC TGGGA TTTCC TGGTC
GGGTC CGTCT TGTCC TGGGT CCCCT GTTGG TCCAG ACCCT AAAGG ACCAG

       2010        2020        2030        2040        2050
CACCA GGAGA GAGAG GACAA CCAGG AGCTC CAGGC CGCCC TGGCG AGAGG
GTGGT CCTCT CTCTC CTGTT GGTCC TCGAG GTCCG GCGGG ACCGC TCTCC
```

```
       2060        2070        2080        2090        2100
GGCTT CGGGG GCTTG AAGGG CCAGA AAGGT GATAC AATTT CTTGT AATGT
CCGAA GCCCC CGAAC TTCCC GGTCT TTCCA CTATG TTAAA GAACA TTACA

       2110        2120        2130        2140        2150
CACCT ACCCT GGGAG GCCAG GACCC CCAGG TTTTG ATGGA CCTCC AGGAC
GTGGA TGGGA CCCTC CGGTC CTGGG GGTCC AAAAC TACCT GGAGG TCCTG

       2160        2170        2180        2190        2200
CAAAG GGATT TCCAG GTCCT CAAGG CGCTC CAGGG TTGAG GTGCT TGGAT
GTTTC CCTAA AGGTC CAGGA GTTCC GCGAG GTCCC AACTC CACGA ACCTA

       2210        2220        2230        2240        2250
GGGCC AAAAG GTCAG CGTGG CAAAC CAGGA ATGTC AGAGA TACCC GGTCC
CCCGG TTTTC CAGTC GCACC GTTTG GTCCT TACAG TCTCT ATGGG CCAGG

       2260        2270        2280        2290        2300
ACCTG GTTTT CGTGG TGACA TGGGA GATCC GGGTT TTGAA GGTGA AAAGG
TGGAC CAAAA GCACC ACTGT ACCCT CTAGG CCCAA AACTT CCACT TTTCC

       2310        2320        2330        2340        2350
GGTCC TCCCT TCTTG GGCCG CCAGG CTTTC CTGGT TCTCG TGGAG CAAAT
CCAGG AGGGA AGAAC CCGGC GGTCC GAAAG GACCA AGAGC ACCTC GTTTA

       2360        2370        2380        2390        2400
GGTCA GAAAG GAATG ATGGG AGACA TTGCC TATGG CCTCC CAGGC CCCCC
CCAGT CTTTC CTTAC TACCC TCTGT AACGG ATACC GGAGG GTCCG GGGGG

       2410        2420        2430        2440        2450
AGGAA AGAGA GGTCC TTCCG GAGTG CCAGG GTCCA AAGGA CACAG AGGTG
TCCTT TCTCT CCAGG AAGGC CTCAC GGTCC CAGGT TTCCT GTGTC TCCAC

       2460        2470        2480        2490        2500
ATTCA GGACG TCCAG GGTTT GCAGG GCCAG CCGGC AAGCC TGGAT CCCCA
TAAGT CCTGC AGGTC CCAAA CGTCC CGGTC GGCCG TTCGG ACCTA GGGGT

       2510        2520        2530        2540        2550
GGTCT CAAAG GTCCC AGAGG CAGAG AGGGA AGTGC TGGGT TTCAA GGTAT
CCAGA GTTTC CAGGG TCTCC GTCTC TCCCT TCACG ACCCA AAGTT CCATA

       2560        2570        2580        2590        2600
CCCGG GTCCA CCTGG CCATT CCTGC AAAAG AGGCG CTCCA GGGAT ACCAG
GGGCC CAGGT GGACC GGTAA GGACG TTTTC TCCGC GAGGT CCCTA TGGTC

       2610        2620        2630        2640        2650
GGCAA CCGGG GCTCC CTGGG GCCCC GGGGA GTCCA GGTGC CCCAG GTTGG
CCGTT GGCCC CGAGG GACCC CGGGG CCCCT CAGGT CCACG GGGTC CAACC

       2660        2670        2680        2690        2700
AAAGG ACAGC GAGGG GATGT GGGGC CTCCG GGTCC AGCTG GAATG AAAGG
TTTCC TGTCG CTCCC CTACA CCCCG GAGGC CCAGG TCGAC CTTAC TTTCC

       2710        2720        2730        2740        2750
CCTCC CTGGA GTCCC TGGAC GGCCA GGGGC AGAGG GACCC CTAGG ATTCC
GGAGG GACCT CAGGG ACCTG CCGGT CCCCG TCTCC CTGGG GATCC TAAGG
```

```
       2760        2770        2780        2790        2800
CAGGA GTCCC AGGCC CCTTA GGGGA TGATG GGCTG CCTGG TCTTC CAGGC
GTCCT CAGGG TCCGG GGAAT CCCCT ACTAC CCGAC GGACC AGAAG GTCCG

       2810        2820        2830        2840        2850
CCGAA GGGAC CCCAG GGGCT GCCTG GCTTC CCCGG TTTTC CAGGG GAGAG
GGCTT CCCTG GGGTC CCCGA CGGAC CGAAG GGGCC AAAAG GTCCC CTCTC

       2860        2870        2880        2890        2900
AGGAA AGCCT GGCCC AGAGG GACGC CCTGG CAGCA AGGGG GAACC AGGAG
TCCTT TCGGA CCGGG TCTCC CTGCG GGACC GTCGT TCCCC CTTGG TCCTC

       2910        2920        2930        2940        2950
AGGAT GGTCG GCCTG GCTTC TTCGG AGACC AAGGG GTGAA AGGTG CCAAA
TCCTA CCAGC CGGAC CGAAG AAGCC TCTGG TTCCC CACTT TCCAC GGTTT

       2960        2970        2980        2990        3000
GGAGC GAGAG GACCC CCAGG AGATG AAGGA GAGAT GGCGG TCTTT TCCCA
CCTCG CTCTC CTGGG GGTCC TCTAC TTCCT CTCTA CCGCC AGAAA AGGGT

       3010        3020        3030        3040        3050
AAAAG GGAAA ACCGG GGAAC CTGGA CCTCC AGGAG ATGAT GGATT CCCAG
TTTTC CCTTT TGGCC CCTTG GACCT GGAGG TCCTC TACTA CCTAA GGGTC

       3060        3070        3080        3090        3100
GAGAA GAAGG TGATA AAGGC AATCC AGGGA TGCCG GGGAG GAGAG GAGAG
CTCTT CTTCC ACTAT TTCCG TTAGG TCCCT ACGGC CCCTC CTCTC CTCTC

       3110        3120        3130        3140        3150
CCTGG AAGAT TCGGA GCACC TGGAT TTCAC AGAGG GGAGC CCGGG AGAAC
GGACC TTCTA AGCCT CGTGG ACCTA AAGTG TCTCC CCTCG GGCCC TCTTG

       3160        3170        3180        3190        3200
CGGGC AGCCA GGGCT TCCTG GCCCC CCAGG CCCCC CAGGC TCACC TGGGC
GCCCG TCGGT CCCGA AGGAC CGGGG GGTCC GGGGG GTCCG AGTGG ACCCG

       3210        3220        3230        3240        3250
TGAGA GGGAT TATTG GTTTT CCGGG ATTTC CAGGT GACCA GGGTG AGCCA
ACTCT CCCTA ATAAC CAAAA GGCCC TAAAG GTCCA CTGGT CCCAC TCGGT

       3260        3270        3280        3290        3300
GGTTC TCCAG GGTCT CCTGG ACTTT CAGGA ATTGA TGGAA TGAGA GGACC
CCAAG AGGTC CCAGA GGACC TGAAA GTCCT TAACT ACCTT ACTCT CCTGG

       3310        3320        3330        3340        3350
TAAAG GAAGC AAAGG TGACC CTGCA AGTCA GTTCG GCCCA CCTGG TCCAA
ATTTC CTTCG TTTCC ACTGG GACGT TCAGT CAAGC CGGGT GGACC AGGTT

       3360        3370        3380        3390        3400
AGGGT GAACC AGGTA GCCCT GGATG TCCAG GACAA GTTGG AGCAC CTGGA
TCCCA CTTGG TCCAT CGGGA CCTAC AGGTC CTGTT CAACC TCGTG GACCT

       3410        3420        3430        3440        3450
GAGCA GGGCT TGCCT GGTGT TCAAG GGCTC AGAGG ACCAC CTGGA AGGCC
CTCGT CCCGA ACGGA CCACA AGTTC CCGAG TCTCC TGGTG GACCT TCCGG
```

```
       3460        3470        3480        3490        3500
AGGAT TACCT GGCTC CTCCG GACCA CCAGG GTGTC CAGGT AATCA AGGGG
TCCTA ATGGA CCGAG GAGGC CTGGT GGTCC CACAG GTCCA TTAGT TCCCC

       3510        3520        3530        3540        3550
TGCCT GGGCT GACAG GGCCT CCAGG AGAAA CAGGG GATCC TGGGC CAAGA
ACGGA CCCGA CTGTC CCGGA GGTCC TCTTT GTCCC CTAGG ACCCG GTTCT

       3560        3570        3580        3590        3600
GGCAT GATGG GAGAT CCAGG GCCAC CAGGT CTTCC AGGAA TAAAA GGTCC
CCGTA CTACC CTCTA GGTCC CGGTG GTCCA GAAGG TCCTT ATTTT CCAGG

       3610        3620        3630        3640        3650
CTCCG GGTCG CCGGG TCTGA ATGGC TTGCA TGGTT TAAAG GGTCA GAAAG
GAGGC CCAGC GGCCC AGACT TACCG AACGT ACCAA ATTTC CCAGT CTTTC

       3660        3670        3680        3690        3700
GAGCC AAAGG TGCTT CAGGT TTGCA CGAAG TGGGC CCACC CGGTC CAGTG
CTCGG TTTCC ACGAA GTCCA AACGT GCTTC ACCCG GGTGG GCCAG GTCAC

       3710        3720        3730        3740        3750
GGCGT ACCTG GGCTG AAAGG GGAGC CGGGA GACCC TGGGA GCCCA GGAAT
CCGCA TGGAC CCGAC TTTCC CCTCG GCCCT CTGGG ACCCT CGGGT CCTTA

       3760        3770        3780        3790        3800
TTCTC CCCCA GGCCT TTCTG GAGAA AGAGG CCCCC CCGGC CCCCC AGGGA
AAGAG GGGGT CCGGA AAGAC CTCTT TCTCC GGGGG GGCCG GGGGG TCCCT

       3810        3820        3830        3840        3850
GACCT GGAGC CCCTG GTCCT GCAGG TGCCA CAGGA AGAGC TGCTG AAGGT
CTGGA CCTCG GGGAC CAGGA CGTCC ACGGT GTCCT TCTCG ACGAC TTCCA

       3860        3870        3880        3890        3900
CACGT TCCTG ACCCA GGTCC ACCCG GAGAT GTGGG TCCTC CTGGC CCCGA
GTGCA AGGAC TGGGT CCAGG TGGGC CTCTA CACCC AGGAG GACCG GGGCT

       3910        3920        3930        3940        3950
TGGTC CGAGA GGAGC GCCCG GGCCC CCAGG CCCCC CTGGG AGCGT TGACC
ACCAG GCTCT CCTCG CGGGC CCGGG GGTCC GGGGG GACCC TCGCA ACTGG

       3960        3970        3980        3990        4000
TTCTG AGAGG GGAAC CAGGA GACTG TGGGC CGCCG GGGCC TCCAG GTCCC
AAGAC TCTCC CCTTG GTCCT CTGAC ACCCG GCGGC CCCGG AGGTC CAGGG

       4010        4020        4030        4040        4050
CCGGG CCCAC CCGGC CCTCC AGGAC ACCAA GGCTT CCCAG GATGC GATGG
GGCCC GGGTG GGCCG GGAGG TCCTG TGGTT CCGAA GGGTC CTACG CTACC

       4060        4070        4080        4090        4100
AAAAG ACGGC CAGAA AGGAC CAATA GGATT CCCGG GGCTG CAGGG GCCTC
TTTTC TGCCG GTCTT TCCTG GTTAT CCTAA GGGCC CCGAC GTCCC CGGAG

       4110        4120        4130        4140        4150
AGGGA CTTCC TGGGC TCCCT GGGGA GAAGG GTCTG CTGGG CATTC CAGGC
TCCCT GAAGG ACCCG AGGGA CCCCT CTTCC CAGAC GACCC GTAAG GTCCG
```

```
       4160        4170        4180        4190        4200
CGGCG AGGGC ACCCC GGTCC TCCAG GTTCC AGAGG TGAAC CAGGG CCTCC
GCCGC TCCCG TGGGG CCAGG AGGTC CAAGG TCTCC ACTTG GTCCC GGAGG

       4210        4220        4230        4240        4250
TGCAG ATGTG GATTC GTGCC CCCGC ATCCC CGGGC TTCCT GGGGT ACCAG
ACGTC TACAC CTAAG CACGG GGGCG TAGGG GCCCG AAGGA CCCCA TGGTC

       4260        4270        4280        4290        4300
GCCCA AGAGG ACCAG AAGGA GCCAT GGGGG TCCCT GGAAG GAGAG GGCCC
CGGGT TCTCC TGGTC TTCCT CGGTA CCCCC AGGGA CCTTC CTCTC CCGGG

       4310        4320        4330        4340        4350
CCAGG ACCAG GGTGC AAAGG AGAGC CGGGA CTGGA GGGCA GGAGG GGCGA
GGTCC TGGTC CCACG TTTCC TCTCG GCCCT GACCT CCCGT CCTCC CCGCT

       4360        4370        4380        4390        4400
GGCCG GCCTC CCAGG ACCTC CGGGG CCTCC CGGAA GCACA GGGGA CGCAG
CCGGC CGGAG GGTCC TGGAG GCCCC GGAGG GCCTT CGTGT CCCCT GCGTC

       4410        4420        4430        4440        4450
GAGAA GCCGG TTGCC CCGGA GCACC AGGGC CCCCC GGGCC CATTG GGGAC
CTCTT CGGCC AACGG GGCCT CGTGG TCCCG GGGGG CCCGG GTAAC CCCTG

       4460        4470        4480        4490        4500
CCCGG GCCCA GAGGG TTTGG GCCTG GATAC CTCAG TGGCT TCCTC CTGGT
GGGCC CGGGT CTCCC AAACC CGGAC CTATG GAGTC ACCGA AGGAG GACCA

       4510        4520        4530        4540        4550
TCTCC ACAGT CAGAC GGATG GGGAG CCCAC CTGCC CCGCG GGCAT GCCCA
AGAGG TGTCA GTCTG CCTAC CCCTC GGGTG GACGG GGCGC CCGTA CGGGT

       4560        4570        4580        4590        4600
GGCCC TGGAC GGGCT ACAGT CTGTT ATACC TGGAA GGACA GGAGA AGGCC
CCGGG ACCTG CCCGA TGTCA GACAA TATGG ACCTT CCTGT CCTCT TCCGG

       4610        4620        4630        4640        4650
CACAA TCAGG ACCTT GGTCT GGCAG GGTCT TGCCT TCCCA TGTTC AGCAC
GTGTT AGTCC TGGAA CCAGA CCGTC CCAGA ACGGA AGGGT ACAAG TCGTG

       4660        4670        4680        4690        4700
CCTGC CCTTT GCCTA CTGCA ACATC CACCA AGTGT GCCAC TACGC CCGGA
GGACG GGAAA CGGAT GACGT TGTAG GTGGT TCACA CGGTG ATGCG GGCCT

       4710        4720        4730        4740        4750
GAAAC GACCG GTCCT ACTGG CTGGC CAGTG CTGCG CCCCT GCCCA TGATG
CTTTG CTGGC CAGGA TGACC GACCG GTCAC GACGC GGGGA CGGGT ACTAC

       4760        4770        4780        4790        4800
CCGCT CTCAG AGGAG GAGAT CCGCC CCTAC ATCAG CCGCT GTGCC GTGTG
GGCGA GAGTC TCCTC CTCTA GGCGG GGATG TAGTC GGCGA CACGG CACAC

       4810        4820        4830        4840        4850
CGAGG CCCCA GCCCC CGTGG TGGCG CTGCA CAGCC AGGAC CGGTC CATCC
GCTCC GGGGT CGGGG GCACC ACCGC GACGT GTCGG TCCTG GCCAG GTAGG
```

```
       4860        4870        4880        4890        4900
CGCCG TGTCC ACGGA GCTGG AGGAG CCTCT GGATC GGGTA CTCGT TCCTG
GCGGC ACAGG TGCCT CGACC TCCTC GGAGA CCTAG CCCAT GAGCA AGGAC

       4910        4920        4930        4940        4950
ATGCA CACAG GGGCT GGAGA CCAAG GAGGA GGGCA GGCCC TCATG TCCCC
TACGT GTGTC CCCGA CCTCT GGTTC CTCCT CCCGT CCGGG AGTAC AGGGG

       4960        4970        4980        4990        5000
CGGCA GCTGT CTGGA AGATT TCCGA GCCGC ACCGT TCCTC GAATG CCAAG
GCCGT CGACA GACCT TCTAA AGGCT CGGCG TGGCA AGGAG CTTAC GGTTC

       5010        5020        5030        5040        5050
GCCGA CAGGG AACTT GCCAC TTTTT TGCAA ATGAG TATAG CTTCT GGCTG
CGGCT GTCCC TTGAA CGGTG AAAAA ACGTT TACTC ATATC GAAGA CCGAC

       5060        5070        5080        5090        5100
ACGAC GGTGA GGCCT GACTT GCAGT TTTCC TCGGC ACCAT CACCA GACAC
TGCTG CCACT CCGGA CTGAA CGTCA AAAGG AGCCG TGGTA GTGGT CTGTG

       5110        5120        5130        5140        5150
CTTGA AAGAA AGCCA GGCCC AGCGC CAGAA AATCA GCAGG TGCCA GGTAT
GAACT TTCTT TCGGT CCGGG TCGCG GTCTT TTAGT CGTCC ACGGT CCATA

       5160
GCATA TGACT AGT
CGTAT ACTGA TCA
```

APPENDIX 5

Canine COL4A4 nucleotide alignment

```
Consensus key (see documentation for details)
* - single, fully conserved residue
  - no consensus


CLUSTAL W (1.81) multiple sequence alignment


Canine_A4_Normal_5-21      CTAGAGGATCTACTAGTCATATGGATGTGAGAGTCTTCTGGAAGAGAAGA
Human_COL4A4               --------------------------------------------------
Mouse_COL4A4               --------------------------------------------------


Canine_A4_Normal_5-21      GAAACCTGACAATCTTTTTTAAGTCCCAAGACTTGACTGTGAGGTAATCC
Human_COL4A4               ----------------------------------------ATGTGGTCT
Mouse_COL4A4               --------------------------------------------------


Canine_A4_Normal_5-21      ATGCACATGGCATCAGTAAAGTGTTCATTCAGGTGGATCAAGCCCTTGGC
Human_COL4A4               CTGCACATAGTACTAATGAGGTGCTCCTTCAGATTGACCAAGTCCTTGGC
Mouse_COL4A4               ---------------ATGAGGTGCTTTTTCAGATGGACCAAGTCCTTTGT
                                         * * *** *  ***** * ** **** **** *


Canine_A4_Normal_5-21      CAGAGATCCCTGGTCACTTATATTTATCCTCTTTTCTGTACAACATGCAT
Human_COL4A4               CACAGGTCCCTGGTCACTTATACTCATTCTCTTTTCTGTACAATATGTAT
Mouse_COL4A4               CACAGCCCCCTGGTCACTAATATTTATCCTCTTTACTATACAATATGAAT
                           ** **  *********** *** * ** ****** ** ***** *** **


Canine_A4_Normal_5-21      ATGGGAGTGGAAAGAAGTTTGTCGGCCCCTGTGGAGGAAGAGATTGCTCG
Human_COL4A4               ATGGGAGTGGAAAGAAATACATTGGTCCTTGTGGAGGAAGAGATTGCTCT
Mouse_COL4A4               ATGGGAGTGGCAAGAAGTATGGCGGCCCCTGCGGAGGAAGAAACTGCTCT
                           ********** ***** *     ** ** ** ********* * *****


Canine_A4_Normal_5-21      GTGTGCCATTGCTTTCCTGAAAAAGGGTCTCGGGGTCAACCAGGACCGCC
Human_COL4A4               GTTTGCCACTGTGTTCCTGAAAAGGGGTCTCGGGGTCCACCAGGACCACC
Mouse_COL4A4               GTCTGCCAGTGTTTTCCTGAGAAGGGATCTCGGGGTCACCCAGGACCACT
                           ** ***** **  ******* ** ** **********  ******** *


Canine_A4_Normal_5-21      GGGGCCACAGGGTCCCATTGGACCCCTGGGACTGCCAGGACCCACAGGAA
Human_COL4A4               AGGGCCACAGGGTCCAATTGGACCCCTGGGAGCCCCAGGACCCATTGGGC
Mouse_COL4A4               GGGACCACAGGGACCAATAGGGCCCTTGGGACCCCTGGGACCTATTGGGA
                            ** ******** ** ** ** *** *****   *  ***** *  **


Canine_A4_Normal_5-21      TTCCAGGAGAGAAAGGGATGAGAGGTGACAGTGGCCCTCCTGGAGCAGCA
Human_COL4A4               TTTCAGGAGAGAAAGGAATGAGAGGGGACCGCGGCCCTCCTGGAGCAGCA
Mouse_COL4A4               TTCCAGGAGAAAAAGGAGAGAGAGGAGACAGTGGCTCACCTGGACCACCA
                           ** ******* *****   ****** *** * *** * ****** ** **


Canine_A4_Normal_5-21      GGTGACAAAGGTGACAAGGGTCCAACTGGTGTTCCTGGATTTCCAGGGTT
Human_COL4A4               GGGGACAAAGGAGATAAGGGTCCAACTGGTGTTCCTGGATTTCCAGGTTT
Mouse_COL4A4               GGAGAGAAAGGAGACAAGGGTCCAACTGGTGTCCCTGGATTTCCAGGTGT
                           ** ** ***** ** ***************** **************  *


Canine_A4_Normal_5-21      GGATGGCATACCTGGCCACCCGGGGCCTCCTGGATCCAGAGGCAAGCCTG
Human_COL4A4               AGATGGCATACCTGGGCACCCAGGGCCTCCTGGACCCAGAGGCAAACCTG
Mouse_COL4A4               GGATGGTGTACCTGGTCACCCAGGGCCTCCTGGACCCAGAGGCAAACCTG
                            *****  ******* ***** ************ ********** ****


Canine_A4_Normal_5-21      GCATGCATGGCTACAATGGTTCACGAGGTGATCCAGGGTTTCCAGGAAAA
Human_COL4A4               GTATGAGTGGCCACAATGGCTCAAGAGGTGACCCAGGGTTTCCAGGAGGA
Mouse_COL4A4               GTGTGGACGGCTACAATGGCTCACGTGGTGATCCGGGCTATCCAGGAGAA
                           *  **   *** ******* *** * ***** ** ** * *******  *
```

```
Canine_A4_Normal_5-21      AGAGGAGTTCCTGGCCCAGGAGGCCCCCCAGGCCTTCCTGGGGAAAGGGG
Human_COL4A4               AGAGGAGCTCTTGGCCCAGGAGGCCCCCTAGGCCATCCTGGGGAAAAGGG
Mouse_COL4A4               AGAGGAGCTCCAGGCCCAGGCGGCCCCCCTGGTCAGCCTGGGGAAAATGG
                           ******* **  ******** *******  ** *  **********  **

Canine_A4_Normal_5-21      AGAAAAAGGAAATTCAGTGTTCATTTTAGGTGCCATTAAAGGTATTCAGG
Human_COL4A4               AGAAAAAGGAAATTCAGTGTTCATTTTAGGTGCCGTTAAAGGTATTCAGG
Mouse_COL4A4               AGAAAAAGGAAGATCTGTGTACATTACTGGTGGCGTTAAAGGTATTCAGG
                           ***********  ** **** ****   **** * ***************

Canine_A4_Normal_5-21      GAGACAGGGGGGACCCAGACCTGCCTGGCTTGCCGGGACCGAAGGGGGCA
Human_COL4A4               GAGACAGAGGGGACCCAGGACTGCCTGGCTTACCAGGATCTTGGGGTGCA
Mouse_COL4A4               GAGACCGTGGGGACCCAGGACCACCCGGCTTACCAGGATCTAGGGGTGCA
                           ***** * **********  *  ** ***** ** *** *   *** ***

Canine_A4_Normal_5-21      GGAGGACCGGCAGGTCCCAAGGGACAGCCAGGAGAGCCTGGGTTAACAGG
Human_COL4A4               GGAGGACCGGCAGGTCCCACAGGATATCCTGGAGAGCCAGGGTTAGTGGG
Mouse_COL4A4               CAAGGGTCACCAGGGCCCATGGGGCATGCAGGCGCACCAGGGCTAGCAGG
                             ***  *  **** ****  **  *  * ** *  ** *** **   **

Canine_A4_Normal_5-21      TGCTCCGGGCCATCCTGGGAGACCAGGCTTGAAGGGTAATCCTGGTGTGG
Human_COL4A4               ACCTCCGGGCCAACCAGGGCGTCCAGGTTTGAAGGGAAATCCCGGTGTGG
Mouse_COL4A4               ACCTATAGGTCATCCTGGAAGCCCTGGGTTGAAGGGAAATCCTGCCACTG
                             **   ** ** ** **  * ** ** ******** ***** *     *

Canine_A4_Normal_5-21      GAGTAAAGGGGCAAATGGGAGACCCGGGTGAGGTTGGCCAGCAGGGTTCT
Human_COL4A4               GAGTAAAGGGGCAAATGGGAGACCCGGGTGAGGTTGGTCAGCAAGGTTCT
Mouse_COL4A4               GACTGAAGGGACAAAGGGGAGAGCCGGGTGAAGTTGGCCAGCGTGGTCCT
                           ** * ***** **** ****** ******** ***** ****  *** **

Canine_A4_Normal_5-21      CCTGGACCCACCTTATTGGTACAGCCACCTGATTCCTGTCTGTATAAAGG
Human_COL4A4               CCTGGACCCACCCTGTTGGTAGAGCCACCTGACTTTTGTCTCTATAAAGG
Mouse_COL4A4               CCTGGGCCCACGCTGTTGGTACAGCCGCCTGATTTGAGTATCTATAAAGG
                           ***** *****  * ****** **** ***** *   ** * ********

Canine_A4_Normal_5-21      AGAAAAGGGCATAAAAGGAATGCCTGGTATGATTGGACCTCCAGGACTAC
Human_COL4A4               AGAAAAGGGTATAAAAGGAATTCCTGGAATGGTTGGACTGCCAGGACCAC
Mouse_COL4A4               AGAAAAGGGTGTAAAAGGAATGCCTGGAATGATTGGTCCCCCAGGACCTC
                           *********  ********** ***** *** **** *  *******  *

Canine_A4_Normal_5-21      CGGGACCCAAGGGAGAACCTGGAATTGGACCAAAAGGAGAGAAGGGTATT
Human_COL4A4               CAGGACGCAAGGGAGAATCTGGTATTGGGGCAAAAGGAGAAAAAGGTATT
Mouse_COL4A4               CAGGTCGCAAGGGAGCGCCTGGTGTCGGGATCAAAGGAGAGAAAGGTATC
                           * ** * ********   ****  * **    ******** ** *****

Canine_A4_Normal_5-21      CCTGGGTTCTCAGGACCTCGGGGTGATCCGGGTTCCTATGGCTCTCCAGG
Human_COL4A4               CCTGGATTTCCAGGGCCTCGGGGGGATCCTGGTTCCTATGGATCTCCAGG
Mouse_COL4A4               CCTGGGTTCCCAGGACCCCGGGGTGAGCCTGGTTCGCATGGGCCTCCAGG
                           ***** **  **** ** ***** ** ** *****  ****  *******

Canine_A4_Normal_5-21      TTTTCCAGGATTAAAGGGGAAACCAGGACTGTTTGGAGATCCTGGATCAT
Human_COL4A4               TTTTCCAGGATTAAAGGGAGAACTAGGACTGGTTGGAGATCCTGGGCTAT
Mouse_COL4A4               TTTTCCAGGATTCAAGGGGATACAAGGAGCAGCTGGAGAGCCTGGGCTGT
                           ************ *****   ** ****     ****** *****    *

Canine_A4_Normal_5-21      TTGGATTTGTTGGCCCAAAGGGGGATCCTGGAGACCGTGGGTACCCAGGA
Human_COL4A4               TTGGATTAATTGGCCCAAAGGGGGATCCTGGAAATCGAGGGCACCCAGGA
Mouse_COL4A4               TCGGATTTCTTGGTCCAAAGGGGGATCTGGGAGATCGCGGGTACCCGGGA
                           * *****  **** *************  *** * ** *** **** ***

Canine_A4_Normal_5-21      CCACCGGGTGTTCTGGTAACTCCATCTGTTCCACTCAAAGGCCCTCCAGG
Human_COL4A4               CCACCAGGTGTTTTGGTGACTCCACCTCTTCCACTTAAAGGCCCACCAGG
Mouse_COL4A4               CCACCAGGCATTTTGTTGACTCCAGCTCCACCACTCAAAGGTGTTCCAGG
                           ***** **  ** ** * ****** **   ***** *****    *****
```

```
Canine_A4_Normal_5-21      GGATCCGGGGCGCCCTGGCCACTATGGAGAAATGGGGTCCGTTGGACCAC
Human_COL4A4               GGACCCAGGGTTCCCTGGCCGCTATGGAGAAACAGGGGATGTTGGACCAC
Mouse_COL4A4               GGACCCAGGACCCCCTGGCTACTATGGGGAGATTGGAGATGTTGGATTAC
                           *** ** **   *******  ****** ** *  **    ******  **

Canine_A4_Normal_5-21      CTGGTCCCCCTGGTCCCCCAGGTCCACCAGGGGAAGACTGTGCAGGCATG
Human_COL4A4               CTGGTCCCCCAGGTCTCTTGGGCAGACCAGGGGAAGCCTGTGCAGGCATG
Mouse_COL4A4               CAGGTCCTCCAGGGCCCCCAGGCAGACCAGGGGAAACTTGTCCAGGCATG
                           * ***** ** ** * *   **   **********   *** ********

Canine_A4_Normal_5-21      ATGGGACCCCCTGGACCAAGAGGATTTCCTGGTCATCCAGGATTTCCAGG
Human_COL4A4               ATAGGACCCCCTGGGCCACAAGGATTTCCTGGTCTTCCTGGGCTTCCAGG
Mouse_COL4A4               ATGGGACCTCCTGGACCACCAGGGGTTCCTGGTCCTCCAGGATTTCCAGG
                           ** ***** ***** ***  ***  ********* *** **  *******

Canine_A4_Normal_5-21      GGCAGCTGGTATCCCAGGGAGAGCGGATTTAAGTCCAGGAAAACCAGGGA
Human_COL4A4               AGAAGCTGGTATTCCTGGGAGACCTGATTCTGCTCCAGGAAAACCAGGGA
Mouse_COL4A4               AGAAGCTGGGGTTCCTGGAAGACTCGACTGTGCCCCAGGAAAACCTGGGA
                            * ******  * ** ** ***   ** *     *********** ****

Canine_A4_Normal_5-21      AACCGGGACCTCCTGGGTTGCCCGGAGCACCAGGGCTGCAAGGACCCCCA
Human_COL4A4               AGCCAGGATCACCTGGCTTGCCTGGAGCACCAGGCCTGCAGGGCCTCCCA
Mouse_COL4A4               AGCCAGGGCTACCCGGCCTGCCTGGCGCACCAGGACCAGAGGGGCCCCCT
                           * ** **    ** **  **** ** ******** *   * ** * ***

Canine_A4_Normal_5-21      GGATCAGATGTTATCTATTGTAGTGTTGGGCATCCTGGACCACAAGGGAT
Human_COL4A4               GGATCAAGTGTGATATACTGTAGTGTTGGGAACCCCGGACCACAAGGAAT
Mouse_COL4A4               GGATCAGATGTGATATACTGTCGCCCTGGGTGCCCTGGGCCAATGGGTGA
                           ******  *** ** ** *** *   ****   ** ** ***   **

Canine_A4_Normal_5-21      AAAAGGCAAAATGGGTCCTCCAGGAAGAAGAGGCTCGAAAGGAGAAAAAG
Human_COL4A4               AAAAGGCAAAGTTGGTCCCCCAGGAGGAAGAGGCCCAAAAGGAGAAAAAG
Mouse_COL4A4               AAAGGGAAAAGTGGGTCCTCCAGGAAGAAGAGGAGCCAAAGGAGCAAAAG
                           *** ** *** * ***** ****** *******  * ******* *****

Canine_A4_Normal_5-21      GAAATGCGGGGCTCTGTGCCTGTGAGCCTGGTCCCATGGGCCCACCAGGC
Human_COL4A4               GAAATGAAGGACTCTGTGCCTGTGAGCCTGGACCCATGGGCCCCCCTGGC
Mouse_COL4A4               GAAACAAGGGACTCTGTACCTGTCCACCTGGTCCCATGGGACCCCCTGGT
                           ****    ** ****** *****   ***** ******** ** ** **

Canine_A4_Normal_5-21      CCTCCAGGACTTCCTGGAAGGCAGGGTAGTAAGGGAGACTTGGGGCTCCC
Human_COL4A4               CCTCCAGGACTTCCTGGGAGGCAGGGGAGTAAGGGAGACTTGGGGCTCCC
Mouse_COL4A4               CCCCCAGGACCTCCTGGAAGACAAGGAAGTAAAGGAGACTTAGGGCTCCC
                           ** ******* ****** ** ** ** ***** ******** ********

Canine_A4_Normal_5-21      TGGGTGGATTGGAGAGAAAGGTCACCCAGGCCCTCCTGGTGCTGAAGGAT
Human_COL4A4               TGGCTGGCTTGGAACAAAAGGTGACCCAGGACCTCCTGGTGCTGAAGGAC
Mouse_COL4A4               TGGTTGGCATGGAGAAAAGGGCGATCCAGGTCAGCCTGGTGCTGAAGGAC
                           *** ***  ****   ** **  * ***** *  ***************

Canine_A4_Normal_5-21      CTCCAGGATCACCAGGAAAACATGGCGCCTCAGGACCACCTGGCAGCAAA
Human_COL4A4               CTCCAGGGCTACCAGGAAAGCATGGTGCCTCTGGACCACCTGGCAACAAA
Mouse_COL4A4               CGCCAGGGCCACCAGGAAGACCTGGTGCCATGGGGCCCCCTGGTCACAAA
                           * *****   ********  * *** ***   ** ** *****   ****

Canine_A4_Normal_5-21      GGAGAAAAGGGCGACAGGGTTGTATCAAGAGTGAAAGGGAAAAAAGGAGA
Human_COL4A4               GGGGCGAAGGGTGACATGGTTGTATCAAGAGTTAAAGGGCACAAAGGAGA
Mouse_COL4A4               GGGGAAAAGGGTGACATGGTCATATCAAGAGTGAAAGGGCAAAAAGGAGA
                           ** *  ***** **** ***  ********** ****** * ********

Canine_A4_Normal_5-21      AAGAGGTCCTGATGGGCTCCCAGGATTTCCAGGACAACAGGGACAAAATG
Human_COL4A4               AAGAGGTCCTGATGGGCCCCCAGGATTTCCAGGGCAGCCAGGATCACATG
Mouse_COL4A4               AAGAGGACTGGATGGGCCACCAGGATTTCCGGGGCCACATGGACAAGATG
                           ****** *  *******  *********** ** *  *  ***  * ***
```

```
Canine_A4_Normal_5-21      GTCGGGATGGACTTCCTGGAAAAAAGGGGGATCCAGGCCCCCCGGGGGGT
Human_COL4A4               GTCGGGATGGACATGCTGGAGAAAAAGGGGATCCAGGACCTCCAGGGGAT
Mouse_COL4A4               GTGGGGATGGACGTCCTGGAGAAAGAGGGGATCCTGGGCCACGAGGGGAC
                           ** ********* * ***** ***  ******** ** ** *  ****

Canine_A4_Normal_5-21      CATGAGGATGCAGCCCCAGGTGCTAAAGGGTCTCCTGGACCACCGGGGCC
Human_COL4A4               CATGAAGATGCGACCCCAGGTGGTAAAGGATTTCCTGGACCTCTGGGCCC
Mouse_COL4A4               CATAAGGATGCAGCCCCGGGTGAGAGAGGGCTTCCCGGACTGCCAGGCCC
                           *** * *****  **** ****  * ***   *** ****  *  ** **

Canine_A4_Normal_5-21      CCCAGGCAGAACAGGACCCAGGGGACAACCAGGTCTGGGATTTCCTGGTC
Human_COL4A4               CCCAGGCAAAGCAGGACCTGTTGGGCCCCCAGGACTGGGATTTCCTGGTC
Mouse_COL4A4               TCCAGGAAGAACAGGACCTGAGGGGCCTCCAGGCCTGGGATTTCCAGGCC
                            ***** * * *******    ** *  ***** ************ ** *

Canine_A4_Normal_5-21      CACCAGGAGAGAGAGGACAACCAGGAGCTCCAGGCCGCCCTGGCGAGAGG
Human_COL4A4               CACCAGGAGAGCGAGGCCACCCAGGAGTTCCAGGCCACCCAGGTGTGAGG
Mouse_COL4A4               CACCAGGACAGCGAGGCCTACCCGGAGAGCCCGGGCGTCCGGGCACCAGA
                           ******** ** **** *  ** ****  ** ** *  ** **    **

Canine_A4_Normal_5-21      GGCTTCGGGGGCTTGAAGGGCCAGAAAGGTGATACAATTTCTTGTAATGT
Human_COL4A4               GGCCCTGATGGCTTAAAGGGTCAGAAAGGTGACACAATTTCTTGCAACGT
Mouse_COL4A4               GGCTTTGATGGCACGAAGGGGCAGAAAGGTGACTCCATTCTGTGTAATGT
                           ***   *  ***   ***** ***********  * ***   ** ** **

Canine_A4_Normal_5-21      CACCTACCCTGGGAGGCCAGGACCCCCAGGTTTTGATGGACCTCCAGGAC
Human_COL4A4               AACCTACCCTGGGAGGCATGGCCCTCCAGGTTTTGATGGACCTCCAGGTC
Mouse_COL4A4               AAGCTACCCGGGAAAACCAGGTCTGCCAGGTCTTGATGGACCTCCAGGTC
                            * ****** ** *  *  ** *  ****** ***************** *

Canine_A4_Normal_5-21      CAAAGGGATTTCCAGGTCCTCAAGGCGCTCCAGGGTTGAGGTGCTTGGAT
Human_COL4A4               CGAAGGGATTTCCAGGTCCCCAAGGTGCCCCTGGGCTGAGTGGTTCAGAT
Mouse_COL4A4               TGAAGGGGTTTCCAGGCCCCCCAGGGGCTCCTGGTATGAGATGTCCAGAT
                             ***** ******** ** * *** ** ** **  ****  *    ***

Canine_A4_Normal_5-21      GGGCCAAAAGGTCAGCGTGGCAAACCAGGAATGTCAGAGATACCCGGTCC
Human_COL4A4               GGGCATAAAGGCAGACCTGGCACACCAGGAACAGCGGAAATACCAGGTCC
Mouse_COL4A4               GGGCAGAAAGGCCAGCGGGGCAAACCGGGAATGTCTGGGATTCCAGGTCC
                           ****  *****    *  **** *** ****   * *  ** ** *****

Canine_A4_Normal_5-21      ACCTGGTTTTCGTGGTGACATGGGAGATCCGGGTTTTGAAGGTGAAAAGG
Human_COL4A4               ACCTGGTTTTCGTGGTGACATGGGAGATCCGGGTTTTGGAGGTGAAAAGG
Mouse_COL4A4               ACCTGGTTTTCGTGGTGACATGGGAGATCCAGGCATCAAAGGTGAAAAGG
                           ****************************** **  *   ***********

Canine_A4_Normal_5-21      GGTCCTCCCTTCTTGGGCCGCCAGGCTTTCCTGGTTCTCGTGGAGCAAAT
Human_COL4A4               GGTCCTCCCCTGTTGGGCCCCCAGGCCCTCCCGGCTCACCAGGAGTGAAT
Mouse_COL4A4               GGACCTCCCCTATTGGACCCCCAGGCCCACCTGGGTCTCCTGGAAAGGAT
                           ** ****** * **** ** ******   ** ** ** *  ***    **

Canine_A4_Normal_5-21      GGTCAGAAAGGAATGATGGGAGACATTGCCTATGGCCTCCCAGGCCCCCC
Human_COL4A4               GGTCAGAAAGGAATCCCGGGAGACCCTGCATTTGGTCACCTGGGACCCCC
Mouse_COL4A4               GGCCAGAAAGGAATCCCAGGAGACCCTGCATTTGGTGACCCAGGACCCCC
                           ** ***********    ******  *** * ***   **  ** *****

Canine_A4_Normal_5-21      AGGAAAGAGAGGTCCTTCCGGAGTGCCAGGGTCCAAAGGACACAGAGGTG
Human_COL4A4               GGGAAAGAGGGGTCTTTCAGGAGTGCCAGGGATAAAAGGACCCAGAGGTG
Mouse_COL4A4               TGGAGAGAGGGGTCTTCCAGGAGCACCAGGCATGAAAGGACAGAAAGGTC
                            *** **** **** * * ****  *****    *******  * ****

Canine_A4_Normal_5-21      ATTCAGGACGTCCAGGGTTTGCAGGGCCAGCCGGCAAGCCTGGATCCCCA
Human_COL4A4               ATCCGGGATGTCCAGGGGCTGAAGGGCCAGCTGGCATTCCTGGATTCCTA
Mouse_COL4A4               ACCCAGGGTGCCCAGGGGCTGGAGGCCCTCCTGGCATCCCTGGATCTCCA
                           *  * **  * ******  ** *** **  * ****  *******  * *
```

```
Canine_A4_Normal_5-21      GGTCTCAAAGGTCCCAGAGGCAGAGAGGGAAGTGCTGGGTTTCAAGGTAT
Human_COL4A4               GGTCTCAAAGGTCCCAAAGGCAGAGAGGGACATGCTGGGTTTCCAGGTGT
Mouse_COL4A4               GGTCTCAAAGGTCCCAAAGGTAGAGAGGGAAGTCGTGGCTTTCCAGGGAT
                           *************** *** *********  *  *** **** ***  *

Canine_A4_Normal_5-21      CCCGGGTCCACCTGGCCATTCCTGCAAAAGAGGCGCTCCAGGGATACCAG
Human_COL4A4               CCCAGGTCCACCTGGCCATTCCTGTGAAAGAGGTGCTCCAGGGATACCAG
Mouse_COL4A4               CCCAGGATCCCCTGGCCATTCCTGTGAAAGAGGTGCTCCCGGGATCCCAG
                           *** **  * **************  ******* ***** ***** ****

Canine_A4_Normal_5-21      GGCAACCGGGGCTCCCTGGGGCCCCGGGGAGTCCAGGTGCCCCAGGTTGG
Human_COL4A4               GGCAACCGGGACTCCCTGGGTATCCAGGTAGCCCAGGTGCTCCAGGTGGG
Mouse_COL4A4               GACAACCGGGACTTCCTGGAACTCCAGGTGATCCAGGGGCCCCAGGTTGG
                           * ******** ** *****    ** **    ***** ** ****** **

Canine_A4_Normal_5-21      AAAGGACAGCGAGGGGATGTGGGGCCTCCGGGTCCAGCTGGAATGAAAGG
Human_COL4A4               AAAGGACAGCCGGGAGATGTGGGGCCTCCCGGGCCAGCTGGAATGAAAGG
Mouse_COL4A4               AAAGGCCAGCCAGGAGACATGGGTCCCTCTGGACCAGCTGGGATGAAGGG
                           ***** ****  ** **  **** **  * ** ******** ***** **

Canine_A4_Normal_5-21      CCTCCCTGGAGTCCCTGGACGGCCAGGGGCAGAGGGACCCCTAGGATTCC
Human_COL4A4               CCTCCCCGGACTCCCAGGACGGCCTGGGGCACATGGTCCCCCAGGCCTCC
Mouse_COL4A4               CCTCCCTGGTCTCCCAGGCCTACCAGGAGCTGATGGTCTCCGAGGGCCTC
                           ****** **  **** ** *  ** ** **  * ** * ** ***    *

Canine_A4_Normal_5-21      CAGGAGTCCCAGGCCCCTTAGGGGATGATGGGCTGCCTGGTCTTCCAGGC
Human_COL4A4               CAGGAATCCCAGGTCCCTTTGGAGATGATGGGCTACCTGGTCCTCCAGGT
Mouse_COL4A4               CTGGGATTCCAGGCCCCAATGGGGAAGATGGTCTACCTGGTCTTCCAGGC
                           * **  * ***** ***   ** ** ***** ** ******* ******

Canine_A4_Normal_5-21      CCGAAGGGACCCCAGGGGCTGCCTGGCTTCCCCGGTTTTCCAGGGGAGAG
Human_COL4A4               CCAAAGGGACCCCGGGGGCTGCCTGGTTTCCCAGGTTTTCCCGGAGAAAG
Mouse_COL4A4               CTGAAGGGACTCCCAGGGCTGCCTGGCTTCCCTGGTTTTCCAGGAGAGAG
                           *  ******* **  *********** ***** ******** ** ** **

Canine_A4_Normal_5-21      AGGAAAGCCTGGCCCAGAGGGACGCCCTGGCAGCAAGGGGGAACCAGGAG
Human_COL4A4               AGGAAAGCCTGGTGCAGAGGGATGTCCTGGCGCAAAGGGAGAACCTGGAG
Mouse_COL4A4               AGGAAAGCCAGGGCCAGATGGAGAACCTGGAAGAAAGGGAGAGGTTGGAG
                           ********* **  **** ***   *****    ***** **    ****

Canine_A4_Normal_5-21      AGGATGGTCGGCCTGGCTTCTTCGGAGACCAAGGGGTGAAAGGTGCCAAA
Human_COL4A4               AGAAGGGCATGTCTGGCCTTCCTGGAGACCGGGGACTGAGAGGGGCCAAA
Mouse_COL4A4               AGAAAGGCTGGCCTGGCTTGAAGGGAGACCTGGGAGAGCGAGGCGCCAAA
                           ** * **   * ***** *    *******  **   *  *** ******

Canine_A4_Normal_5-21      GGAGCGAGAGGACCCCCAGGAGATGAAGGAGAGATGGCGGTCTTTTCCCA
Human_COL4A4               GGAGCCATAGGACCTCCCGGAGATGAAGGAGAAATGGCTATCATTTCACA
Mouse_COL4A4               GGAGACAGAGGACTTCCTGGTGATGCAGGAGAA---GCAGTCACTTCCAG
                           ****  * *****  ** ** **** ******    **  **  ***

Canine_A4_Normal_5-21      AAAAGGGAAAACCGGGGAACCTGGACCTCCAGGAGATGATGGATTCCCAG
Human_COL4A4               AAAGGGAACACCTGGGGAACCTGGACCTCCTGGAGATGATGGATTCCCAG
Mouse_COL4A4               AAAGGGTGAACCTGGGGATGCTGGGCCTCCAGGAGATGGTGGGTTCTCAG
                           *** **   * * *****  **** ***** ******* *** *** ***

Canine_A4_Normal_5-21      GAGAAGAAGGTGATAAAGGCAATCCAGGGATGCCGGGGAGGAGAGGAGAG
Human_COL4A4               GAGAAAGAGGTGATAAAGGAACTCCCGGGATGCAAGGGAGAAGAGGAGAG
Mouse_COL4A4               GAGAGAGAGGTGATAAAGGAAGCTCAGGGATGCGAGGGGGAAGAGGGGAT
                           ****   ************ *   * *******  *** * ***** **

Canine_A4_Normal_5-21      CCTGGAAGATTCGGAGCACCTGGATTTCACAGAGGGGAGCCCGGGAGAAC
Human_COL4A4               CTGGGAAGATACGGACCACCTGGATTTCACAGAGGGGAACCTGGTGAGAA
Mouse_COL4A4               CCCGGAAGAGATGGACTACCGGGACTCCACAGAGGGCAGCCTGGGATAGA
                           *  ******   ***  *** *** * ********* * ** **
```

```
Canine_A4_Normal_5-21      CGGGCAGCCAGGGCTTCCTGGCCCCCCAGGCCCCCCAGGCTCACCTGGGC
Human_COL4A4               AGGTCAGCCAGGGCCTCCTGGACCCCCAGGCCCTCCAGGCTCAACTGGTC
Mouse_COL4A4               TGGGCCTCCTGGGCCTCCTGGGCCCCCTGGCCCTCCAGGATCACCTGGAC
                            ** *  ** **** ****** ***** ***** ***** *** **** *

Canine_A4_Normal_5-21      TGAGAGGGATTATTGGTTTTCCGGGATTTCCAGGTGACCAGGGTGAGCCA
Human_COL4A4               TAAGAGGGTTCATTGGTTTTCCAGGACTTCCAGGTGACCAGGGTGAGCCA
Mouse_COL4A4               TGAGAGGAGTCATTGGTTTCCCAGGATTTCCAGGTGACCAGGGTGATCCA
                           * *****  * ******** ** *** ****************** ***

Canine_A4_Normal_5-21      GGTTCTCCAGGGTCTCCTGGACTTTCAGGAATTGATGGAATGAGAGGACC
Human_COL4A4               GGTTCTCCAGGTCCCCCTGGATTTTCAGGAATTGATGGAGCAAGAGGACC
Mouse_COL4A4               GGTTCTCCAGGACCCCCTGGATTCCCAGGAGATGATGGAGCAAGAGGACC
                           ***********  * ****** *  *****  *******   ********

Canine_A4_Normal_5-21      TAAAGGAAGCAAAGGTGACCCTGCAAGTCAGTTCGGCCCACCTGGTCCAA
Human_COL4A4               TAAAGGAAACAAAGGTGACCCTGCCAGTCACTTTGGTCCACCTGGTCCAA
Mouse_COL4A4               TAAAGGATACAAAGGTGACCCTGCCAGTCAGTGTGGTCCACCTGGTCCAA
                           *******  *************** ***** *  ** *************

Canine_A4_Normal_5-21      AGGGTGAACCAGGTAGCCCTGGATGTCCAGGACAAGTTGGAGCACCTGGA
Human_COL4A4               AGGGTGAGCCAGGTAGCCCTGGATGTCCAGGGCATTTTGGAGCATCCGGA
Mouse_COL4A4               AGGGTGAGCCAGGTAGTCCTGGATACCAAGGACGTACTGGAGTCCCCGGA
                           ******* ******** *******  * *** *    *****   * ***

Canine_A4_Normal_5-21      GAGCAGGGCTTGCCTGGTGTTCAAGGGCTCAGAGGACCACCTGGAAGGCC
Human_COL4A4               GAGCAGGGCTTGCCTGGTATTCAAGGGCCCAGAGGATCACCTGGAAGGCC
Mouse_COL4A4               GAGAAAGGCTTTCCTGGAGATGAAGGACCCCGAGGACCACCAGGCAGACC
                           *** * ***** *****   * **** * * ***** **** ** ** **

Canine_A4_Normal_5-21      AGGATTACCTGGCTCCTCCGGACCACCAGGGTGTCCAGGTAATCAAGGGG
Human_COL4A4               AGGGCCACCTGGCTCCTCTGGACCACCAGGGTGCCCAGGTGATCACGGGA
Mouse_COL4A4               TGGACAGCCTGGATCTTTTGGACCACCAGGGTGTCCAGGTGACCCAGGGA
                            **    ***** ** *  ************** ****** * *  ***

Canine_A4_Normal_5-21      TGCCTGGGCTGACAGGGCCTCCAGGAGAAACAGGGGATCCTGGGCCAAGA
Human_COL4A4               TGCCTGGGCTGAGGGGACAGCCAGGAGAAATGGGAGACCCTGGGCCAAGA
Mouse_COL4A4               TGCCCGGACTCAAGGGACATCCAGGAGAAGTGGGAGACCCTGGGCCAAGA
                           **** ** ** *  ** *  *********   ** ** ************

Canine_A4_Normal_5-21      GGCATGATGGGAGATCCAGGGCCACCAGGTCTTCCAGGAATAAAAGGTCC
Human_COL4A4               GGCCTCCAGGGGGATCCAGGGATACCAGGTCCTCCGGGAATAAAAGGTCC
Mouse_COL4A4               GGTGATGCGGGAGATTTCGGGAGGCCAGGTCCTGCTGGAGTGAAAGGGCC
                           **       *** ***   ***   ******* * * *** * ***** **

Canine_A4_Normal_5-21      CTCCGGGTCGCCGGGTCTGAATGGCTTGCATGGTTTAAAGGGTCAGAAAG
Human_COL4A4               CTCCGGATCACCTGGCCTGAACGGCTTGCATGGATTGAAAGGTCAGAAAG
Mouse_COL4A4               TCTCGGGTCTCCTGGTCTGAATGGCTTACATGGTCTGAAGGGTGAGAAAG
                              *** ** ** ** ***** ***** *****  * ** *** ******

Canine_A4_Normal_5-21      GAGCCAAAGGTGCTTCAGGTTTGCACGAAGTGGGCCCACCCGGTCCAGTG
Human_COL4A4               GAACTAAAGGTGCTTCAGGTTTGCATGATGTGGGGCCACCTGGTCCAGTG
Mouse_COL4A4               GAACAAAAGGCGCCTCAGGTTTGCTCGAAATGGGTCCACCTGGGCCAATG
                           ** * ***** ** **********  **  **** ***** ** *** **

Canine_A4_Normal_5-21      GGCGTACCTGGGCTGAAAGGGGAGCCGGGAGACCCTGGGAGCCCAGGAAT
Human_COL4A4               GGAATACCTGGGCTAAAAGGGGAGAGAGGAGACCCTGGGAGCCCAGGAAT
Mouse_COL4A4               GGGATGCCTGGGCAAAAAGGAGAGAAAGGAGACCCTGGGAGTCCAGGGAT
                           **  * *******  ***** ***   ************** ***** **

Canine_A4_Normal_5-21      TTCTCCCCCAGGCCTTTCTGGAGAAAGAGGCCCCCCCGGCCCCCCAGGGA
Human_COL4A4               CTCTCCTCCAGGTCCTCGTGGAAAGAAAGGTCCCCCAGGACCCCCAGGGA
Mouse_COL4A4               TTCTCCTCCAGGTCTTCCTGGAGAAAAGGGCTTCCCAGGACCCCCAGGGA
                            ***** ***** * *  **** * *  **   *** ** **********
```

```
Canine_A4_Normal_5-21      GACCTGGAGCCCCTGGTCCTGCAGGTGCCACAGGAAGAGCTGCTGAAGGT
Human_COL4A4               GTTCAGGACCACCTGGTCCTGCAGGTGCCACAGGAAGAGCTCCTAAGG--
Mouse_COL4A4               GACCAGGACCACCTGGTCCCGCAGGTGCCCCAGGAAGAGCTGCTAAAGGT
                           *  * *** * ******** ********* *********** ** * *

Canine_A4_Normal_5-21      CACGTTCCTGACCCAGGTCCACCCGGAGATGTGGGTCCTCCTGGCCCCGA
Human_COL4A4               -ACATTCCTGACCCGGGTCCACCTGGAGATCAGGGACCTCCTGGTCCTGA
Mouse_COL4A4               GACATTCCTGATCCAGGTCCACCTGGAGACCGGGGACCTCCTGGCCCCGA
                            ** ******* ** ******** *****   *** ******** ** **

Canine_A4_Normal_5-21      TGGTCCGAGAGGAGCGCCCGGGCCCCCAGGCCCCCCTGGGAGCGTTGACC
Human_COL4A4               TGGCCCAAGAGGAGCACCTGGGCCTCCAGGCCTCCCTGGGAGTGTTGACC
Mouse_COL4A4               TGGCCCAAGAGGAGTACCTGGGCCTCCGGGTTCCCCTGGGAATGTTGACC
                           *** ** *******  ** ***** ** **   ********  *******

Canine_A4_Normal_5-21      TTCTGAGAGGGGAACCAGGAGACTGTGGGCCGCCGGGGCCTCCAGGTCCC
Human_COL4A4               TTCTGAGAGGGGAGCCAGGTGACTGTGGTCTACCAGGGCCACCAGGTCCC
Mouse_COL4A4               TTCTGAAAGGGGACCCAGGTGACTGTGGTTTGCCAGGACCACCTGGCTCC
                           ****** ****** ***** ********    ** ** ** ** **  **

Canine_A4_Normal_5-21      CCGGGCCCACCCGGCCCTCCAGGACACCAAGGCTTCCCAGGATGCGATGG
Human_COL4A4               CCTGGCCCACCAGGCCCTCCAGGATACAAAGGCTTTCCAGGATGTGATGG
Mouse_COL4A4               CGAGGCCCACCAGGCCCTCCAGGGTGTCAGGGTCCCCCAGGATGTGATGG
                           *  ******** ***********     * **    ******** *****

Canine_A4_Normal_5-21      AAAAGACGGCCAGAAAGGACCAATAGGATTCCCGGGGCTGCAGGGGCCTC
Human_COL4A4               AAAAGATGGCCAGAAAGGACCAATGGGATTCCCGGGACCGCAGGGACCAC
Mouse_COL4A4               CAAAGATGGCCAGAAAGGACCAATGGGACTTCCAGGGTTGCCAGGGCCAC
                            ***** ***************** *** * ** **   **  ** ** *

Canine_A4_Normal_5-21      AGGGACTTCCTGGGCTCCCTGGGGAGAAGGGTCTGCTGGGCATTCCAGGC
Human_COL4A4               ATGGATTTCCTGGGCCACCTGGAGAGAAGGGTTTACCTGGACCTCCAGGG
Mouse_COL4A4               CTGGTCTTCCTGGGGCACCTGGCGAGAAGGGATTACCTGGCCCTCCAGGC
                             **  ********   ***** ********  * *  **   ******

Canine_A4_Normal_5-21      CGGCGAGGGCACCCCGGTCCTCCAGGTTCCAGAGGTGAACCAGGGCCTCC
Human_COL4A4               AGAAAAGGGCCCACTGGTCTTCCAGGTCCCAGAGGTGAACCGGGGCCACC
Mouse_COL4A4               AGAAAAGGGCCAGTAGGTCCTCCAGGCTGCAGAGGTGAACCTGGGCCTCC
                            *   *****     **** ******   ************ ***** **

Canine_A4_Normal_5-21      TGCAGATGTGGATTCGTGCCCCCGCATCCCCGGGCTTCCTGGGGTACCAG
Human_COL4A4               TGCAGATGTGGATGACTGTCCCCGAATCCCAGGCCTTCCTGGGGCGCCAG
Mouse_COL4A4               TGCAGACGTGGATTCCTGTCCTCGAATCCCAGGACTTCCTGGAGTGCCAG
                           ****** ******   ** ** ** ***** ** ******** *  ****

Canine_A4_Normal_5-21      GCCCAAGAGGACCAGAAGGAGCCATGGGGGTCCCTGGAAGGAGAGGGCCC
Human_COL4A4               GCATGAGAGGACCAGAAGGAGCCATGGGGCTCCCTGGAATGAGAGGCCCC
Mouse_COL4A4               GTCCCAGAGGACCAGAAGGAGCTATGGGGGAACCTGGACGGAGAGGCCTG
                           *    ***************** ******   ******  ****** *

Canine_A4_Normal_5-21      CCAGGACCAGGGTGCAAAGGAGAGCCGGGACTGGAGGGCAGGAGGGGCGA
Human_COL4A4               CCAGGACCAGGGTGCAAAGGAGAGCCTGGGCTGGATGGCAGGAGGGGTGT
Mouse_COL4A4               CCTGGACCAGGGTGCAAAGGAGAGCCTGGACCGGATGGCAGACGGGGCCA
                           ** **********************  ** * *** *****  ****

Canine_A4_Normal_5-21      GGCCGGCCTCCCAGGACCTCCGGGGCCTCCCGGAAGCACAGGGGACGCAG
Human_COL4A4               GGATGGCGTCCCTGGGTCTCCTGGGCCTCCCGGACGTAAAGGTGACACAG
Mouse_COL4A4               GGATGGCATCCCTGGATCTCCCGGGCCTCCTGGACGCAAAGGTGACACTG
                           **  *** **** **  **** ******** *** * * *** *** * *

Canine_A4_Normal_5-21      GAGAAGCCGGTTGCCCCGGAGCACCAGGGCCCCCCGGGCCCATTGGGGAC
Human_COL4A4               GAGAAGACGGCTACCCTGGAGGACCAGGGCCTCCTGGTCCCATTGGGGAT
Mouse_COL4A4               GAGAGGCAGGGTGTCCTGGAGCACCAGGCCCTCCTGGTCCAACTGGTGAT
                           **** *  ** *  ** **** ****** ** ** ** ** * *** **
```

```
Canine_A4_Normal_5-21      CCCGGGCCCAGAGGGTTTGGGCCTGGATACCTCAGTGGCTTCCTCCTGGT
Human_COL4A4               CCTGGGCCCAAAGGGTTTGGCCCTGGATACCTCGGTGGCTTCCTCCTGGT
Mouse_COL4A4               CCTGGGCCCAAGGGGTTTGGCCCTGGATCGCTCAGTGGCTTCCTCCTGGT
                           ** *******  ******** *******  *** ****************

Canine_A4_Normal_5-21      TCTCCACAGTCAGACGGATGGGGAGCCCACCTGCCCCGCGGGCATGCCCA
Human_COL4A4               TCTCCACAGTCAGACGGACCAGGAGCCCACCTGCCCCCTGGGCATGCCCA
Mouse_COL4A4               TCTCCACAGTCAGACTGACCAGGAACCGGCCTGCCCCGTGGGCATGCCTC
                           *************** **   *** **  ********  *********

Canine_A4_Normal_5-21      GGCCCTGGACGGGCTACAGTCTGTTATACCTGGAAGGACAGGAGAAGGCC
Human_COL4A4               GGCTCTGGACTGGGTATAGTCTGTTATACCTGGAAGGGCAAGAGAAAGCT
Mouse_COL4A4               GGCTTTGGACAGGGTACAGCTTGTTGTACATGGAAGGACAGGAGAAGGCG
                           ***  ***** ** ** **  **** *** ******* ** ***** **

Canine_A4_Normal_5-21      CACAATCAGGACCTTGGTCTGGCAGGGTCTTGCCTTCCCATGTTCAGCAC
Human_COL4A4               CACAATCAAGACCTTGGTCTGGCAGGGTCTTGCCTTCCCGTATTTAGCAC
Mouse_COL4A4               CACAATCAAGATCTTGGTTTGGCAGGCTCTTGTCTTCCTGTGTTTAGCAC
                           ******** ** ****** ******* ***** *****  * ** *****

Canine_A4_Normal_5-21      CCTGCCCTTTGCCTACTGCAACATCCACCAAGTGTGCCACTACGCCCGGA
Human_COL4A4               GCTGCCCTTTGCCTACTGCAACATCCACCAGGTGTGCCACTATGCCCAGA
Mouse_COL4A4               TCTGCCCTTTGCCTACTGCAACATCCACCAAGTGTGCCACTACGCCCAGA
                            ***************************** *********** **** **

Canine_A4_Normal_5-21      GAAACGACCGGTCCTACTGGCTGGCCAGTGCTGCGCCCCTGCCCATGATG
Human_COL4A4               GAAACGACAGATCCTACTGGCTGGCCAGCGCTGCGCCCCTCCCCATGATG
Mouse_COL4A4               GAAATGACAGGTCCTACTGGCTGTCCAGTGCGGCTCCTCTTCCTATGATG
                           **** *** * ************ **** ** ** ** ** ** ******

Canine_A4_Normal_5-21      CCGCTCTCAGAGGAGGAGATCCGCCCCTACATCAGCCGCTGTGCCGTGTG
Human_COL4A4               CCACTCTCTGAAGAGGCGATCCGCCCCTATGTCAGCCGCTGTGCGGTATG
Mouse_COL4A4               CCACTCTCAGAGGAAGAGATCCGCTCTTACATCAGCCGCTGTGCGGTGTG
                           ** ***** ** ** * ******* * **  ************** ** **

Canine_A4_Normal_5-21      CGAGGCCCCAGCCCCCGTGGTGGCGCTGCACAGCCAGGACCGGTCCATCC
Human_COL4A4               CGAGGCCCCGGCCCAGGCGGTGGCGGTGCACAGCCAGGACCAGTCCATCC
Mouse_COL4A4               TGAGGCCCCAGCACAGGCCGTGGCGGTCCACAGTCAGGACCAATCCATCC
                            ******** ** *  *  ****** * ***** *******  *******

Canine_A4_Normal_5-21      CGCCGTGTCCACGGAGCTGGAGGAGCCTCTGGATCGGGTACTCGTTCCTG
Human_COL4A4               CCCCATGTCCGCAGACCTGGAGGAGCCTCTGGATCGGGTATTCATTCCTG
Mouse_COL4A4               CACCGTGTCCACGGACATGGAGGAGTCTCTGGATTGGGTACTCATTTCTG
                           * ** ***** * **  ******** ******** ***** ** ** ***

Canine_A4_Normal_5-21      ATGCACACAGGGGCTGGAGACCAAGGAGGAGGGCAGGCCCTCATGTCCCC
Human_COL4A4               ATGCACACAGGAGCTGGGGACCAAGGAGGAGGGCAGGCCCTTATGTCACC
Mouse_COL4A4               ATGCACACTGGAGCTGGGGACCAAGGAGGTGGGCAGGCACTGATGTCACC
                           ******** ** ***** *********** ******** ** ***** **

Canine_A4_Normal_5-21      CGGCAGCTGTCTGGAAGATTTCCGAGCCGCACCGTTCCTCGAATGCCAAG
Human_COL4A4               TGGCAGCTGCCTGGAAGATTTCAGAGCAGCACCATTCCTTGAATGCCAGG
Mouse_COL4A4               TGGCAGCTGCCTAGAGGACTTCAGAGCTGCACCGTTCGTTGAATGCCAGG
                            ******** ** ** ** *** **** ***** *** * ******** *

Canine_A4_Normal_5-21      GCCGACAGGGAACTTGCCACTTTTTTGCAAATGAGTATAGCTTCTGGCTG
Human_COL4A4               GCCGGCAGGGAACTTGCCACTTTTTCGCAAATAAGTATAGCTTCTGGCTC
Mouse_COL4A4               GCCGACAGGGAACCTGTCACTTTTTTGCAAACGAGTACAGCTTCTGGCTG
                           **** ******** ** ******** *****  **** ***********

Canine_A4_Normal_5-21      ACGACGGTGAGGCCTGACTTGCAGTTTTCCTCGGCACCATCACCAGACAC
Human_COL4A4               ACAACGGTGAAAGCAGACTTGCAGTTTTCCTCTGCTCCAGCACCAGACAC
Mouse_COL4A4               ACCACCGTGAACCCAGACCTGCAGTTTGCCTCTGGCCCGTCACCAGACAC
                           ** ** ****   * *** ******** **** *  **  **********
```

```
Canine_A4_Normal_5-21      CTTGAAAGAAAGCCAGGCCCAGCGCCAGAAAATCAGCAGGTGCCAGGTAT
Human_COL4A4               CTTAAAAGAAAGCCAGGCCCAACGCCAGAAAATCAGCCGGTGCCAGGTCT
Mouse_COL4A4               CTTGAAAGAAGTTCAGGCCCAGCGCCGGAAAATCAGCAGGTGCCAGGTTT
                           *** ******   ******** **** ********** ********** *

Canine_A4_Normal_5-21      GCAT-----ATGACTAGT
Human_COL4A4               GCGTGAAGTATAGCTAG-
Mouse_COL4A4               GCATGAAGCACAGCTAA-
                           ** *     *   ***
```

APPENDIX 6

Canine COL4A4 amino acid sequence alignments

```
Consensus key (see documentation for details)
* - single, fully conserved residue
: - conservation of strong groups
. - conservation of weak groups
  - no consensus


CLUSTAL W (1.81) multiple sequence alignment


Canine_A4_Normal_5-21_Transl      RIYSYGCESLLEEKRNLTIFFKSQDLTVRSMHMASVKCSFRWIKPLARDP
Human_COL4A4_Translated_-_Fr      ---------------------------MWSLHIVLMRCSFRLTKSLATGP
Mouse_COL4A4_Translated_-_Fr      ----------------------------------MRCFFRWTKSFVTAP
                                                                   ::* **  *.:.  *

Canine_A4_Normal_5-21_Transl      WSLIFILFSVQHAYGSGKKFVGPCGGRDCSVCHCFPEKGSRGQPGPPGPQ
Human_COL4A4_Translated_-_Fr      WSLILILFSVQYVYGSGKKYIGPCGGRDCSVCHCVPEKGSRGPPGPPGPQ
Mouse_COL4A4_Translated_-_Fr      WSLIFILFTIQYEYGSGKKYGGPCGGRNCSVCQCFPEKGSRGHPGPLGPQ
                                  ****:***::*: ******: ******:****:*.******* *** ***

Canine_A4_Normal_5-21_Transl      GPIGPLGLPGPTGIPGEKGMRGDSGPPGAAGDKGDKGPTGVPGFPGLDGI
Human_COL4A4_Translated_-_Fr      GPIGPLGAPGPIGLSGEKGMRGDRGPPGAAGDKGDKGPTGVPGFPGLDGI
Mouse_COL4A4_Translated_-_Fr      GPIGPLGPLGPIGIPGEKGERGDSGSPGPPGEKGDKGPTGVPGFPGVDGV
                                  *******  ** *:.**** *** *.**..*:**************:**:

Canine_A4_Normal_5-21_Transl      PGHPGPPGSRGKPGMHGYNGSRGDPGFPGKRGVPGPGGPPGLPGERGEKG
Human_COL4A4_Translated_-_Fr      PGHPGPPGPRGKPGMSGHNGSRGDPGFPGGRGALGPGGPLGHPGEKGEKG
Mouse_COL4A4_Translated_-_Fr      PGHPGPPGPRGKPGVDGYNGSRGDPGYPGERGAPGPGGPPGQPGENGEKG
                                  ********.*****: *:********:** **. ***** * ***.****

Canine_A4_Normal_5-21_Transl      NSVFILGAIKGIQGDRGDPDLPGLPGPKGAGGPAGPKGQPGEPGLTGAPG
Human_COL4A4_Translated_-_Fr      NSVFILGAVKGIQGDRGDPGLPGLPGSWGAGGPAGPTGYPGEPGLVGPPG
Mouse_COL4A4_Translated_-_Fr      RSVYITGGVKGIQGDRGDPGPPGLPGSRGAQGSPGPMGHAGAPGLAGPIG
                                  .**:* *.:**********. *****. ** *..** * .* ***.*. *

Canine_A4_Normal_5-21_Transl      HPGRPGLKGNPGVGVKGQMGDPGEVGQQGSPGPTLLVQPPDSCLYKGEKG
Human_COL4A4_Translated_-_Fr      QPGRPGLKGNPGVGVKGQMGDPGEVGQQGSPGPTLLVEPPDFCLYKGEKG
Mouse_COL4A4_Translated_-_Fr      HPGSPGLKGNPATGLKGQRGEPGEVGQRGPPGPTLLVQPPDLSIYKGEKG
                                  :** *******..*:*** *:******:*.*******:*** .:******

Canine_A4_Normal_5-21_Transl      IKGMPGMIGPPGLPGPKGEPGIGPKGEKGIPGFSGPRGDPGSYGSPGFPG
Human_COL4A4_Translated_-_Fr      IKGIPGMVGLPGPPGRKGESGIGAKGEKGIPGFPGPRGDPGSYGSPGFPG
Mouse_COL4A4_Translated_-_Fr      VKGMPGMIGPPGPPGRKGAPGVGIKGEKGIPGFPGPRGEPGSHGPPGFPG
                                  :**:***:* ** ** ** .*:* *********.****:***:*.*****

Canine_A4_Normal_5-21_Transl      LKGKPGLFGDPGSFGFVGPKGDPGDRGYPGPPGVLVTPSVPLKGPPGDPG
Human_COL4A4_Translated_-_Fr      LKGELGLVGDPGLFGLIGPKGDPGNRGHPGPPGVLVTPPLPLKGPPGDPG
Mouse_COL4A4_Translated_-_Fr      FKGIQGAAGEPGLFGFLGPKGDLGDRGYPGPPGILLTPAPPLKGVPGDPG
                                  :**  *  *:** **::***** *:**:*****:*:**. **** *****

Canine_A4_Normal_5-21_Transl      RPGHYGEMGSVGPPGPPGPPGPPGEDCAGMMGPPGPRGFPGHPGFPGAAG
Human_COL4A4_Translated_-_Fr      FPGRYGETGDVGPPGPPGLLGRPGEACAGMIGPPGPQGFPGLPGLPGEAG
Mouse_COL4A4_Translated_-_Fr      PPGYYGEIGDVGLPGPPGPPGRPGETCPGMMGPPGPPGVPGPPGFPGEAG
                                   ** *** *.** *****  * *** *.**:***** *.** **:** **

Canine_A4_Normal_5-21_Transl      IPGRADLSPGKPGKPGPPGLPGAPGLQGPPGSDVIYCSVGHPGPQGIKGK
Human_COL4A4_Translated_-_Fr      IPGRPDSAPGKPGKPGSPGLPGAPGLQGLPGSSVIYCSVGNPGPQGIKGK
Mouse_COL4A4_Translated_-_Fr      VPGRLDCAPGKPGKPGLPGLPGAPGPEGPPGSDVIYCRPGCPGPMGEKGK
                                  :*** * :******** ******** :* ***.****  * *** * ***
```

```
Canine_A4_Normal_5-21_Transl      MGPPGRRGSKGEKGNAGLCACEPGPMGPPGPPGLPGRQGSKGDLGLPGWI
Human_COL4A4_Translated_-_Fr      VGPPGGRGPKGEKGNEGLCACEPGPMGPPGPPGLPGRQGSKGDLGLPGWL
Mouse_COL4A4_Translated_-_Fr      VGPPGRRGAKGAKGNKGLCTCPPGPMGPPGPPGPPGRQGSKGDLGLPGWH
                                  :**** **.** *** ***:* ********** **************

Canine_A4_Normal_5-21_Transl      GEKGHPGPPGAEGSPGSPGKHGASGPPGSKGEKGDRVVSRVKGKKGERGP
Human_COL4A4_Translated_-_Fr      GTKGDPGPPGAEGPPGLPGKHGASGPPGNKGAKGDMVVSRVKGHKGERGP
Mouse_COL4A4_Translated_-_Fr      GEKGDPGQPGAEGPPGPPGRPGAMGPPGHKGEKGDMVISRVKGQKGERGL
                                  * **.** *****.** **: ** **** ** *** *:*****:*****

Canine_A4_Normal_5-21_Transl      DGLPGFPGQQGQNGRDGLPGKKGDPGPPGGHEDAAPGAKGSPGPPGPPGR
Human_COL4A4_Translated_-_Fr      DGPPGFPGQPGSHGRDGHAGEKGDPGPPGDHEDATPGGKGFPGPLGPPGK
Mouse_COL4A4_Translated_-_Fr      DGPPGFPGPHGQDGGDGRPGERGDPGPRGDHKDAPGERGLPGLPGPPGR
                                  ** *****  *..* ** .*::***** *.*:**:** :* **  ****:

Canine_A4_Normal_5-21_Transl      TGPRGQPGLGFPGPPGERGQPGAPGRPGERGFGGLKGQKGDTISCNVTYP
Human_COL4A4_Translated_-_Fr      AGPVGPPGLGFPGPPGERGHPGVPGHPGVRGPDGLKGQKGDTISCNVTYP
Mouse_COL4A4_Translated_-_Fr      TGPEGPPGLGFPGPPGQRGLPGEPGRPGTRGFDGTKGQKGDSILCNVSYP
                                  :** * **********:** ** **:** ** .* ******:* ***:**

Canine_A4_Normal_5-21_Transl      GRPGPPGFDGPPGPKGFPGPQGAPGLRCLDGPKGQRGKPGMSEIPGPPGF
Human_COL4A4_Translated_-_Fr      GRHGPPGFDGPPGPKGFPGPQGAPGLSGSDGHKGRPGTPGTAEIPGPPGF
Mouse_COL4A4_Translated_-_Fr      GKPGLPGLDGPPGLKGFPGPPGAPGMRCPDGQKGQRGKPGMSGIPGPPGF
                                  *: * **:***** ****** ****:   ** **: *.** : *******

Canine_A4_Normal_5-21_Transl      RGDMGDPGFEGEKGSSLLGPPGFPGSRGANGQKGMMGDIAYGLPGPPGKR
Human_COL4A4_Translated_-_Fr      RGDMGDPGFGGEKGSSPVGPPGPPGSPGVNGQKGIPGDPAFGHLGPPGKR
Mouse_COL4A4_Translated_-_Fr      RGDMGDPGIKGEKGTSPIGPPGPPGSPGKDGQKGIPGDPAFGDPGPPGER
                                  ********: ****:* :**** *** * :****: ** *:*  ****:*

Canine_A4_Normal_5-21_Transl      GPSGVPGSKGHRGDSGRPGFAGPAGKPGSPGLKGPRGREGSAGFQGIPGP
Human_COL4A4_Translated_-_Fr      GLSGVPGIKGPRGDPGCPGAEGPAGIPGFLGLKGPKGREGHAGFPGVPGP
Mouse_COL4A4_Translated_-_Fr      GLPGAPGMKGQKGHPGCPGAGGPPGIPGSPGLKGPKGREGSRGFPGIPGS
                                  * .*.** ** :*..* **  **.* **  *****:****  ** *:**.

Canine_A4_Normal_5-21_Transl      PGHSCKRGAPGIPGQPGLPGAPGSPGAPGWKGQRGDVGPPGPAGMKGLPG
Human_COL4A4_Translated_-_Fr      PGHSCERGAPGIPGQPGLPGYPGSPGAPGGKGQPGDVGPPGPAGMKGLPG
Mouse_COL4A4_Translated_-_Fr      PGHSCERGAPGIPGQPGLPGTPGDPGAPGWKGQPGDMGPSGPAGMKGLPG
                                  *****:************** **.***** *** **:**.**********

Canine_A4_Normal_5-21_Transl      VPGRPGAEGPLGFPGVPGPLGDDGLPGLPGPKGPQGLPGFPGFPGERGKP
Human_COL4A4_Translated_-_Fr      LPGRPGAHGPPGLPGIPGPFGDDGLPGPPGPKGPRGLPGFPGFPGERGKP
Mouse_COL4A4_Translated_-_Fr      LPGLPGADGLRGPPGIPGPNGEDGLPGLPGLKGLPGLPGFPGFPGERGKP
                                  :** ***.*  * **:*** *:***** ** **  ***************

Canine_A4_Normal_5-21_Transl      GPEGRPGSKGEPGEDGRPGFFGDQGVKGAKGARGPPGDEGEMAVFSQKGK
Human_COL4A4_Translated_-_Fr      GAEGCPGAKGEPGEKGMSGLPGDRGLRGAKGAIGPPGDEGEMAIISQKGT
Mouse_COL4A4_Translated_-_Fr      GPDGEPGRKGEVGEKGWPGLKGDLGERGAKGDRGLPGDAG-EAVTSRKGE
                                  *.:* ** *** **.* .*: ** * :****  * *** *  *: *:**

Canine_A4_Normal_5-21_Transl      TGEPGPPGDDGFPGEEGDKGNPGMPGRRGEPGRFGAPGFHRGEPGRTGQP
Human_COL4A4_Translated_-_Fr      PGEPGPPGDDGFPGERGDKGTPGMQGRRGELGRYGPPGFHRGEPGEKGQP
Mouse_COL4A4_Translated_-_Fr      PGDAGPPGDGGFSGERGDKGSSGMRGGRGDPGRDGLPGLHRGQPGIDGPP
                                  .*:.*****.**.**.****..** * **: ** * **:***:**  * *

Canine_A4_Normal_5-21_Transl      GLPGPPGPPGSPGLRGIIGFPGFPGDQGEPGSPGSPGLSGIDGMRGPKGS
Human_COL4A4_Translated_-_Fr      GPPGPPGPPGSTGLRGFIGFPGLPGDQGEPGSPGPPGFSGIDGARGPKGN
Mouse_COL4A4_Translated_-_Fr      GPPGPPGPPGSPGLRGVIGFPGFPGDQGDPGSPGPPGFPGDDGARGPKGY
                                  * *********.****.*****:*****:*****.**:.* ** *****

Canine_A4_Normal_5-21_Transl      KGDPASQFGPPGPKGEPGSPGCPGQVGAPGEQGLPGVQGLRGPPGRPGLP
Human_COL4A4_Translated_-_Fr      KGDPASHFGPPGPKGEPGSPGCPGHFGASGEQGLPGIQGPRGSPGRPGPP
Mouse_COL4A4_Translated_-_Fr      KGDPASQCGPPGPKGEPGSPGYQGRTGVPGEKGFPGDEGPRGPPGRPGQP
                                  ******: *************  *: *..**:*:** :* **.***** *
```

```
Canine_A4_Normal_5-21_Transl      GSSGPPGCPGNQGVPGLTGPPGETGDPGPRGMMGDPGPPGLPGIKGPSGS
Human_COL4A4_Translated_-_Fr      GSSGPPGCPGDHGMPGLRGQPGEMGDPGPRGLQGDPGIPGPPGIKGPSGS
Mouse_COL4A4_Translated_-_Fr      GSFGPPGCPGDPGMPGLKGHPGEVGDPGPRGDAGDFGRPGPAGVKGPLGS
                                  ** *******: *:*** * *** *******  ** * ** .*:*** **

Canine_A4_Normal_5-21_Transl      PGLNGLHGLKGQKGAKGASGLHEVGPPGPVGVPGLKGEPGDPGSPGISPP
Human_COL4A4_Translated_-_Fr      PGLNGLHGLKGQKGTKGASGLHDVGPPGPVGIPGLKGERGDPGSPGISPP
Mouse_COL4A4_Translated_-_Fr      PGLNGLHGLKGEKGTKGASGLLEMGPPGPMGMPGQKGEKGDPGSPGISPP
                                  ***********:**:****** ::*****:*:** *** ***********

Canine_A4_Normal_5-21_Transl      GLSGERGPPGPPGRPGAPGPAGATGRAAEGHVPDPGPPGDVGPPGPDGPR
Human_COL4A4_Translated_-_Fr      GPRGKKGPPGPPGSSGPPGPAGATGRAPK-DIPDPGPPGDQGPPGPDGPR
Mouse_COL4A4_Translated_-_Fr      GLPGEKGFPGPPGRPGPPGPAGAPGRAAKGDIPDPGPPGDRGPPGPDGPR
                                  *  *::* ***** .*.******.***.: .:******** *********

Canine_A4_Normal_5-21_Transl      GAPGPPGPPGSVDLLRGEPGDCGPPGPPGPPGPPGPPGPPGHQGFPGCDGKDG
Human_COL4A4_Translated_-_Fr      GAPGPPGLPGSVDLLRGEPGDCGLPGPPGPPGPPGPPGYKGFPGCDGKDG
Mouse_COL4A4_Translated_-_Fr      GVPGPPGSPGNVDLLKGDPGDCGLPGPPGSRGPPGPPGCQGPPGCDGKDG
                                  *.***** **.****:*:***** *****. ******* :* ********

Canine_A4_Normal_5-21_Transl      QKGPIGFPGLQGPQGLPGLPGEKGLLGIPGRRGHPGPPGSRGEPGPPADV
Human_COL4A4_Translated_-_Fr      QKGPMGFPGPQGPHGFPGPPGEKGLPGPPGRKGPTGLPGPRGEPGPPADV
Mouse_COL4A4_Translated_-_Fr      QKGPMGLPGLPGPPGLPGAPGEKGLPGPPGRKGPVGPPGCRGEPGPPADV
                                  ****:*:**  ** *:** ****** * ***:*  * ** **********

Canine_A4_Normal_5-21_Transl      DSCPRIPGLPGVPGPRGPEGAMGVPGRRGPPGPGCKGEPGLEGRRGEAGL
Human_COL4A4_Translated_-_Fr      DDCPRIPGLPGAPGMRGPEGAMGLPGMRGPPGPGCKGEPGLDGRRGVDGV
Mouse_COL4A4_Translated_-_Fr      DSCPRIPGLPGVPGPRGPEGAMGEPGRRGLPGPGCKGEPGPDGRRGQDGI
                                  *.*********.** ******** ** ** ********** :****  *:

Canine_A4_Normal_5-21_Transl      PGPPGPPGSTGDAGEAGCPGAPGPPGPIGDPGPRGFGPGYLSGFLLVLHS
Human_COL4A4_Translated_-_Fr      PGSPGPPGRKGDTGEDGYPGGPGPPGPIGDPGPKGFGPGYLGGFLLVLHS
Mouse_COL4A4_Translated_-_Fr      PGSPGPPGRKGDTGEAGCPGAPGPPGPTGDPGPKGFGPGSLSGFLLVLHS
                                  **.***** .**:** * **.****** *****:***** *.********

Canine_A4_Normal_5-21_Transl      QTDGEPTCPAGMPRPWTGYSLLYLEGQEKAHNQDLGLAGSCLPMFSTLPF
Human_COL4A4_Translated_-_Fr      QTDQEPTCPLGMPRLWTGYSLLYLEGQEKAHNQDLGLAGSCLPVFSTLPF
Mouse_COL4A4_Translated_-_Fr      QTDQEPACPVGMPRLWTGYSLLYMEGQEKAHNQDLGLAGSCLPVFSTLPF
                                  *** **:** **** ********:******************:******

Canine_A4_Normal_5-21_Transl      AYCNIHQVCHYARRNDRSYWLASAAPLPMMPLSEEEIRPYISRCAVCEAP
Human_COL4A4_Translated_-_Fr      AYCNIHQVCHYAQRNDRSYWLASAAPLPMMPLSEEAIRPYVSRCAVCEAP
Mouse_COL4A4_Translated_-_Fr      AYCNIHQVCHYAQRNDRSYWLSSAAPLPMMPLSEEEIRSYISRCAVCEAP
                                  ************:********:************* **.*:*********

Canine_A4_Normal_5-21_Transl      APVVALHSQDRSIPPCPRSWRSLWIGYSFLMHTGAGDQGGGQALMSPGSC
Human_COL4A4_Translated_-_Fr      AQAVAVHSQDQSIPPCPQTWRSLWIGYSFLMHTGAGDQGGGQALMSPGSC
Mouse_COL4A4_Translated_-_Fr      AQAVAVHSQDQSIPPCPRTWRSLWIGYSFLMHTGAGDQGGGQALMSPGSC
                                  * .**:****:******::*****************************

Canine_A4_Normal_5-21_Transl      LEDFRAAPFLECQGRQGTCHFFANEYSFWLTTVRPDLQFSSAPSPDTLKE
Human_COL4A4_Translated_-_Fr      LEDFRAAPFLECQGRQGTCHFFANKYSFWLTTVKADLQFSSAPAPDTLKE
Mouse_COL4A4_Translated_-_Fr      LEDFRAAPFVECQGRQGTCHFFANEYSFWLTTVNPDLQFASGPSPDTLKE
                                  *********:**************:********..****:*.*:******

Canine_A4_Normal_5-21_Transl      SQAQRQKISRCQVCIL--
Human_COL4A4_Translated_-_Fr      SQAQRQKISRCQVCVKYS
Mouse_COL4A4_Translated_-_Fr      VQAQRRKISRCQVCMKHS
                                   ****:********:
```

APPENDIX 7

Comparison of ECS-HN and normal COL4A3

```
          30        40        50        60        70        80
ECS_A3 CGATGGTGTACCAGGATGCAATGGTTCTAAGGGTGAGCAAGGGTTTCCAGGCCTTCCAGG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
NORMAL CGATGGTGTACCAGGATGCAATGGTTCTAAGGGTGAGCAAGGGTTTCCAGGCCTTCCAGG
         30        40        50        60        70        80

        90       100       110       120       130       140
ECS_A3 GACACCAGGCTACCCAGGGATCCCGGGTGCTGTTGGCCTGAAAGGAGAAAAAGGTGCTCC
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
NORMAL GACACCAGGCTACCCAGGGATCCCGGGTGCTGTTGGCCTGAAAGGAGAAAAAGGTGCTCC
         90       100       110       120       130       140

      150       160       170       180       190       200
ECS_A3 TGCCACAGAAGGTATAGAACTTGATGGAAGAGGTGACCCTGGGTTGCCAGGAGCTCCAGG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
NORMAL TGCCACAGAAGGTATAGAACTTGATGGAAGAGGTGACCCTGGGTTGCCAGGAGCTCCAGG
        150       160       170       180       190       200

      210       220       230       240       250       260
ECS_A3 ATTCCAGGGTTTACCRGGCCTTCCAGGCTTTCCGGGACCTGCTGGTCCACCTGGCCCTCC
       ::::::::::::::.::::::::::::::::::::::::::::::::::::::::::::
NORMAL ATTCCAGGGTTTACCAGGCCTTCCAGGCTTTCCGGGACCTGCTGGTCCACCTGGCCCTCC
        210       220       230       240       250       260

      270       280       290       300       310       320
ECS_A3 GGGATTCTTAGGCTTTCCAGGAGCCATGGGACCTCCAGGACCTAAGGGTCAAATGGGCGA
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
NORMAL GGGATTCTTAGGCTTTCCAGGAGCCATGGGACCTCCAGGACCTAAGGGTCAAATGGGCGA
        270       280       290       300       310       320

      330       340       350       360       370       380
ECS_A3 TAAAGTGATAGGACAAAAAGGAGAGCGGGGTGTGAAAGGATTAACAGGACCTCCCGGACC
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
NORMAL TAAAGTGATAGGACAAAAAGGAGAGCGGGGTGTGAAAGGATTAACAGGACCTCCCGGACC
        330       340       350       360       370       380

      390       400       410       420       430       440
ECS_A3 ACCAGGAACAGTTATTGTGACGCTAACCGGCCCTGATAACAGAACGGACCTYAAGGGGGA
       ::::::::::::::::::::::::::::::::::::::::::::::::::::.::::::::
NORMAL ACCAGGAACAGTTATTGTGACGCTAACCGGCCCTGATAACAGAACGGACCTCAAGGGGGA
        390       400       410       420       430       440

      450       460       470       480       490       500
ECS_A3 GAAGGGAGACAAGGGAGCCATGGGGAAGCCTGGACCTCCTGGGCCCTTGGGACCGCCTGG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
NORMAL GAAGGGAGACAAGGGAGCCATGGGGAAGCCTGGACCTCCTGGGCCCTTGGGACCGCCTGG
        450       460       470       480       490       500

      510       520       530       540       550       560
ECS_A3 AGAATCCTATGGATCTGAAAAAGGTGCTCCTGGAGAACCTGGCCCGCAGGGAAAACCTGG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
NORMAL AGAATCCTATGGATCTGAAAAAGGTGCTCCTGGAGAACCTGGCCCGCAGGGAAAACCTGG
        510       520       530       540       550       560

      570       580       590       600       610       620
ECS_A3 CAAAGATGGTGCCCCTGGTTTCCCTGGCACTGAGGGGCCCAAAGGCAACAGAGGCTTCCC
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
NORMAL CAAAGATGGTGCCCCTGGTTTCCCTGGCACTGAGGGGCCCAAAGGCAACAGAGGCTTCCC
        570       580       590       600       610       620
```

```
      630       640       650       660       670       680
ECS_A3 TGGGTTACGGGGTGAAGACGGCATTAAGGGGTGGAAAGGGGACATTGGCCCGCCAGGATT
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
NORMAL TGGGTTACGGGGTGAAGACGGCATTAAGGGGTGGAAAGGGGACATTGGCCCGCCAGGATT
        630       640       650       660       670       680

      690       700       710       720       730       740
ECS_A3 TCGTGGTCCAACAGAATATTATGATGCATACCAGGAAAAGGGTGATGAAGGAATTCCAGG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
NORMAL TCGTGGTCCAACAGAATATTATGATGCATACCAGGAAAAGGGTGATGAAGGAATTCCAGG
        690       700       710       720       730       740

      750       760       770       780       790       800
ECS_A3 CCCACCAGGCCCCAAAGGAGCTCGTGGCCCCCAGGGTCCCAGCGGTCCCCCTGGAGTTCT
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
NORMAL CCCACCAGGCCCCAAAGGAGCTCGTGGCCCCCAGGGTCCCAGCGGTCCCCCTGGAGTTCT
        750       760       770       780       790       800

      810       820       830       840       850       860
ECS_A3 TGGAAGTGCTGGGTCATCAAAGCCTGGCCTCAGAGGAGCTCCTGGATCTCCAGGCATGAA
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
NORMAL TGGAAGTGCTGGGTCATCAAAGCCTGGCCTCAGAGGAGCTCCTGGATCTCCAGGCATGAA
        810       820       830       840       850       860

      870       880       890       900       910       920
ECS_A3 AGGACGTAAAGGGGAACAAGGGCCCCCGGGAAAGAATGCAGTGGGGCCTCCTGGGTCCCC
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
NORMAL AGGACGTAAAGGGGAACAAGGGCCCCCGGGAAAGAATGCAGTGGGGCCTCCTGGGTCCCC
        870       880       890       900       910       920

      930       940       950       960       970       980
ECS_A3 AGGTTGTCCTGGTTCACCAGGCCCCATAGGGCTGCCGGGATATCCAGGACCACCAGGTGA
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
NORMAL AGGTTGTCCTGGTTCACCAGGCCCCATAGGGCTGCCGGGATATCCAGGACCACCAGGTGA
        930       940       950       960       970       980

      990      1000      1010      1020      1030      1040
ECS_A3 CATCGTGTTTCGCAAAGGTCCACCTGGGGTTGGTGGACTCCCCGGCCATGTTGGGTTTCC
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
NORMAL CATCGTGTTTCGCAAAGGTCCACCTGGGGTTGGTGGACTCCCCGGCCATGTTGGGTTTCC
        990      1000      1010      1020      1030      1040

     1050      1060      1070      1080      1090      1100
ECS_A3 AGGCATCCCAGGAGTCGAGGGGCCCAAAGGGGAACCAGGCCTCTTGTGCGTACAGTGTGC
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
NORMAL AGGCATCCCAGGAGTCGAGGGGCCCAAAGGGGAACCAGGCCTCTTGTGCGTACAGTGTGC
       1050      1060      1070      1080      1090      1100

     1110      1120      1130      1140      1150      1160
ECS_A3 TTATATCCCAGGGCGTCCAGGTCTCCCAGGACTGCCAGGGTTGGATGGCATAAAAGGAAT
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
NORMAL TTATATCCCAGGGCGTCCAGGTCTCCCAGGACTGCCAGGGTTGGATGGCATAAAAGGAAT
       1110      1120      1130      1140      1150      1160

     1170      1180      1190      1200      1210      1220
ECS_A3 CCCAGGAGGACAAGGGGCAGCTGGCATTAAAGGAAGCCCAGGGTCCCCAGGAAGTGCGGG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
NORMAL CCCAGGAGGACAAGGGGCAGCTGGCATTAAAGGAAGCCCAGGGTCCCCAGGAAGTGCGGG
       1170      1180      1190      1200      1210      1220

     1230      1240      1250      1260      1270      1280
ECS_A3 TCTTCCAGGATTTCCAGGATTCCCGGGTGCTCCGGGTGCTCCAGGACTTAAAGGAGAAAA
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
NORMAL TCTTCCAGGATTTCCAGGATTCCCGGGTGCTCCGGGTGCTCCAGGACTTAAAGGAGAAAA
       1230      1240      1250      1260      1270      1280
```

```
     1290      1300      1310      1320      1330      1340
ECS_A3 AGGAGAAACATCTGAGCCTGAGGGACAGGTGGGTGCTCCAGGGGACCCCGGGCTCAGAGG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
NORMAL AGGAGAAACATCTGAGCCTGAGGGACAGGTGGGTGCTCCAGGGGACCCCGGGCTCAGAGG
       1290      1300      1310      1320      1330      1340

     1350      1360      1370      1380      1390      1400
ECS_A3 CCATCCTGGAAGAAAGGGCTTGGATGGAATTCCTGGAACTCCTGGAATTAAAGGACCACC
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
NORMAL CCATCCTGGAAGAAAGGGCTTGGATGGAATTCCTGGAACTCCTGGAATTAAAGGACCACC
       1350      1360      1370      1380      1390      1400

     1410      1420      1430      1440      1450      1460
ECS_A3 AGGACCCAAAGGTGAACCGGCCCTGAGTGGTGAGAAGGGGGACCAGGGTCCTCCAGGGGA
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
NORMAL AGGACCCAAAGGTGAACCGGCCCTGAGTGGTGAGAAGGGGGACCAGGGTCCTCCAGGGGA
       1410      1420      1430      1440      1450      1460

     1470      1480      1490      1500      1510      1520
ECS_A3 TCCTGGGATCCCAGGGTCCCCAGGACCTGCAGGACCACCTGGACCACCAAACTATGGACC
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
NORMAL TCCTGGGATCCCAGGGTCCCCAGGACCTGCAGGACCACCTGGACCACCAAACTATGGACC
       1470      1480      1490      1500      1510      1520

     1530      1540      1550      1560      1570      1580
ECS_A3 ACAGGGAGAGCCTGGTCCAAAGGGCACCCAAGGAGTTCCTGGAGCCCCTGGACCACCTGG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
NORMAL ACAGGGAGAGCCTGGTCCAAAGGGCACCCAAGGAGTTCCTGGAGCCCCTGGACCACCTGG
       1530      1540      1550      1560      1570      1580

     1590      1600      1610      1620      1630      1640
ECS_A3 AGAAGCCGGTCCTAAGGGAGAATTCGGTATTTCGACACCAGTCCCAGGGCCCCCAGGACC
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
NORMAL AGAAGCCGGTCCTAAGGGAGAATTCGGTATTTCGACACCAGTCCCAGGGCCCCCAGGACC
       1590      1600      1610      1620      1630      1640

     1650      1660      1670      1680      1690      1700
ECS_A3 TCCAGGGCCCCCTGGTTATGCTGGCCCCCGAGGTCCACCTGGTATCCCTGGATCCATAGG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
NORMAL TCCAGGGCCCCCTGGTTATGCTGGCCCCCGAGGTCCACCTGGTATCCCTGGATCCATAGG
       1650      1660      1670      1680      1690      1700

     1710      1720      1730      1740      1750      1760
ECS_A3 AAAATGTGGTGATCCGGGTCTTCCTGGGCCTGATGGTGAACCAGGAATTCCAGGAATTGG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
NORMAL AAAATGTGGTGATCCGGGTCTTCCTGGGCCTGATGGTGAACCAGGAATTCCAGGAATTGG
       1710      1720      1730      1740      1750      1760

     1770      1780      1790      1800      1810      1820
ECS_A3 CTTCCCTGGGCCCCCTGGACCTAAGGGAGACCAAGGTTTTCCAGGAGCAAAAGGAGCACC
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
NORMAL CTTCCCTGGGCCCCCTGGACCTAAGGGAGACCAAGGTTTTCCAGGAGCAAAAGGAGCACC
       1770      1780      1790      1800      1810      1820

     1830      1840      1850      1860      1870      1880
ECS_A3 GGGTTGTCCAGGAGAAATGGGGAAGCCCGGGTCACCTGGAGAACCAGGTCTCCCAGGAGC
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
NORMAL GGGTTGTCCAGGAGAAATGGGGAAGCCCGGGTCACCTGGAGAACCAGGTCTCCCAGGAGC
       1830      1840      1850      1860      1870      1880

     1890      1900      1910      1920      1930      1940
ECS_A3 CAAGGGAGAACCAGGACTAGCCATACCTGGAGAACCAGGAATACCAGGTTTTCCAGGAGA
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
NORMAL CAAGGGAGAACCAGGACTAGCCATACCTGGAGAACCAGGAATACCAGGTTTTCCAGGAGA
       1890      1900      1910      1920      1930      1940
```

```
     1950      1960      1970      1980      1990      2000
ECS_A3 AAGAGGCAATTCTGGGGAAAATGGAGAAATTGGACTCCCTGGACTTCCAGGTCTCCCTGG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
NORMAL AAGAGGCAATTCTGGGGAAAATGGAGAAATTGGACTCCCTGGACTTCCAGGTCTCCCTGG
       1950      1960      1970      1980      1990      2000

     2010      2020      2030      2040      2050      2060
ECS_A3 AATTCCAGGAACTGGAGGGCTTGATGGACCACCAGGGGATCCAGGGAAGCCTGGACCACC
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
NORMAL AATTCCAGGAACTGGAGGGCTTGATGGACCACCAGGGGATCCAGGGAAGCCTGGACCACC
       2010      2020      2030      2040      2050      2060

     2070      2080      2090      2100      2110      2120
ECS_A3 TGGAGAAAAAGGACCCCCAGGAAGTTGCACAGAGGGTCCCAGAGGAGCCCACGGACTTCC
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
NORMAL TGGAGAAAAAGGACCCCCAGGAAGTTGCACAGAGGGTCCCAGAGGAGCCCACGGACTTCC
       2070      2080      2090      2100      2110      2120

     2130      2140      2150      2160      2170      2180
ECS_A3 AGGATTAAATGGATTGGAAGGGCAACCAGGCAGAAGAGGTGAAACAGGGCCAAAAGGAGA
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
NORMAL AGGATTAAATGGATTGGAAGGGCAACCAGGCAGAAGAGGTGAAACAGGGCCAAAAGGAGA
       2130      2140      2150      2160      2170      2180

     2190      2200      2210      2220      2230      2240
ECS_A3 CCCAGGTATTCCAGGCTTGGATAGATCAGGCTTTCCTGGAGAACCTGGACCACCAGGAAT
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Normal CCCAGGTATTCCAGGCTTGGATAGATCAGGCTTTCCTGGAGAACCTGGACCACCAGGAAT
       2190      2200      2210      2220      2230      2240

     2250      2260      2270      2280      2290      2300
ECS_A3 GCCAGGTCATCGAGGTGAGATAGGACCACCTGGCCAAAAAGGATATCCAGGAAATCCGGG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Normal GCCAGGTCATCGAGGTGAGATAGGACCACCTGGCCAAAAAGGATATCCAGGAAATCCGGG
       2250      2260      2270      2280      2290      2300

     2310      2320      2330      2340      2350      2360
ECS_A3 ATTTTTAGGACTACCAGGTGAAGATGGARTGATTGGGATGATGGGCTATCCAGGAAACAC
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Normal ATTTTTAGGACTACCAGGTGAAGATGGAGTGATTGGGATGATGGGCTATCCAGGAAACAC
       2310      2320      2330      2340      2350      2360

     2370      2380      2390      2400      2410      2420
ECS_A3 TGGCCCTCCTGGGCCTCCCGGGATCCCAGGCACCCCAGGACAGAGGGGTAGCTTTGGAAT
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Normal TGGCCCTCCTGGGCCTCCCGGGATCCCAGGCACCCCAGGACAGAGGGGTAGCTTTGGAAT
       2370      2380      2390      2400      2410      2420

     2430      2440      2450      2460      2470      2480
ECS_A3 TCCAGGTGCAAAGGGTGAGAAAGGGCCCCCAGGAGCCAAGGGGGAAGAAGGAGAGAAAGG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Normal TCCAGGTGCAAAGGGTGAGAAAGGGCCCCCAGGAGCCAAGGGGGAAGAAGGAGAGAAAGG
       2430      2440      2450      2460      2470      2480

     2490      2500      2510      2520      2530      2540
ECS_A3 ACCTATGGGGCCTTGTCGAATAACCAGCTCAGTGGGGGACAAAGGAGAACCAGGCCTCAA
       ::::::::::::::::: ::::::::::::::::::::::::::::::::::::::::::
Normal ACCTATGGGGCCTTGTCAAATAACCAGCTCAGTGGGGGACAAAGGAGAACCAGGCCTCAA
       2490      2500      2510      2520      2530      2540

     2550      2560      2570      2580      2590      2600
ECS_A3 AGGGTTTGCAGGAAAGCCGGGTGAGAAAGGAAACAGAGGCATTCCAGGGTTACCAGGTTT
       ::::::::::::::::::::::: ::::::::::::::::::::::::::::::::::::
Normal AGGGTTTGCAGGAAAGCCGGGTGAAAAAGGAAACAGAGGCATTCCAGGGTTACCAGGTTT
       2550      2560      2570      2580      2590      2600
```

```
     2610      2620      2630      2640      2650      2660
ECS_A3 GAAAGGATTCGAAGGGCCACCTGGACCACCAGGCCCACCAGGCCCCCGAGGAGATCCAGG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Normal GAAAGGATTCGAAGGGCCACCTGGACCACCAGGCCCACCAGGCCCCCGAGGAGATCCAGG
       2610      2620      2630      2640      2650      2660

     2670      2680      2690      2700      2710      2720
ECS_A3 CAGCATTGGGAATCCTGGAGAAGCAGGACCACGTGGAAGTCCAGGAAACATGGGGAACAT
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Normal CAGCATTGGGAATCCTGGAGAAGCAGGACCACGTGGAAGTCCAGGAAACATGGGGAACAT
       2670      2680      2690      2700      2710      2720

     2730      2740      2750      2760      2770      2780
ECS_A3 GGGGATGCCAGGTCTTAAAGGACACAGGGGACCTTTGGGACTACCAGGTGTCACTGGAAG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Normal GGGGATGCCAGGTCTTAAAGGACACAGGGGACCTTTGGGACTACCAGGTGTCACTGGAAG
       2730      2740      2750      2760      2770      2780

     2790      2800      2810      2820      2830      2840
ECS_A3 ACCAGGCCTCCCAGGTGTTCACGGTCTCCAAGGAGACAAGGGAGAGCCAGGTTATTCAGC
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Normal ACCAGGCCTCCCAGGTGTTCACGGTCTCCAAGGAGACAAGGGAGAGCCAGGTTATTCAGC
       2790      2800      2810      2820      2830      2840

     2850      2860      2870      2880      2890      2900
ECS_A3 AGGTACAAGGCCAGGACCACCGGGACTAAAGGGAGATCCAGGATTACCAGGTGCCATGGG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Normal AGGTACAAGGCCAGGACCACCGGGACTAAAGGGAGATCCAGGATTACCAGGTGCCATGGG
       2850      2860      2870      2880      2890      2900

     2910      2920      2930      2940      2950      2960
ECS_A3 AAGGAAAGGAGAAAGAGGGCTACCTGGCACCCCTGGCCATTCGGGGCCTGCTGGAACTGA
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Normal AAGGAAAGGAGAAAGAGGGCTACCTGGCACCCCTGGCCATTCGGGGCCTGCTGGAACTGA
       2910      2920      2930      2940      2950      2960

     2970      2980      2990      3000      3010      3020
ECS_A3 GGGAGCCCCTGGACATCCCGGAAGTCCTGGCCACCCAGGAAAGCCAGGCCCTGATGGAGA
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Normal GGGAGCCCCTGGACATCCCGGAAGTCCTGGCCACCCAGGAAAGCCAGGCCCTGATGGAGA
       2970      2980      2990      3000      3010      3020

     3030      3040      3050      3060      3070      3080
ECS_A3 TTTGGGGTTAAAAGGCATCAAAGGCTTCCCTGGGTCCCCAGGAGTCAAAGGACCTCCAGG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Normal TTTGGGGTTAAAAGGCATCAAAGGCTTCCCTGGGTCCCCAGGAGTCAAAGGACCTCCAGG
       3030      3040      3050      3060      3070      3080

     3090      3100      3110      3120      3130      3140
ECS_A3 ACCTCCAGGAATCCTAGGATCTCCTGGACCTATGGGCATGAGAGGCAACCAGGGCCGTGA
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Normal ACCTCCAGGAATCCTAGGATCTCCTGGACCTATGGGCATGAGAGGCAACCAGGGCCGTGA
       3090      3100      3110      3120      3130      3140

     3150      3160      3170      3180      3190      3200
ECS_A3 TGGAATCCCTGGTCCCCCAGGAGAAAAGGGAGAAACAGGTTTGCTGGGGGCACATCCAGG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Normal TGGAATCCCTGGTCCCCCAGGAGAAAAGGGAGAAACAGGTTTGCTGGGGGCACATCCAGG
       3150      3160      3170      3180      3190      3200

     3210      3220      3230      3240      3250      3260
ECS_A3 CCCAGGAGGGATCCCTGGTCCTCCAGGAGCCAAAGGAGACAAGGGAGCCCCAGGCCTACC
       :::::::::: :::::::::::::::::::::::::::: ::: ::::::::::::::
Normal CCCAGGAGGGAGCCCTGGTCCTCCAGGAGCCAAAGGAGACAGGGGCGCCCCAGGCCTACC
       3210      3220      3230      3240      3250      3260
```

```
     3270      3280      3290      3300      3310      3320
ECS_A3 AGGCCCCCCAGGCAGGAAAGGGGCAGTGGGAGATGCCGGGCCACGGGGACCCACTGGCAT
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Normal AGGCCCCCCAGGCAGGAAAGGGGCAGTGGGAGATGCCGGGCCACGGGGACCCACTGGCAT
       3270      3280      3290      3300      3310      3320

     3330      3340      3350      3360      3370      3380
ECS_A3 GACGGGACTCCCAGGGCCACCAGGTTTTCCTGGTGCAGTCATCCCTGGCCAGAAAGGAAA
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Normal GACGGGACTCCCAGGGCCACCAGGTTTTCCTGGTGCAGTCATCCCTGGCCAGAAAGGAAA
       3330      3340      3350      3360      3370      3380

     3390      3400      3410      3420      3430      3440
ECS_A3 TCAAGGTCCACCAGGCTTCAGAGGAAACCCAGGTGAGCCTGGTCCTCCAGGCCCTCCAGG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Normal TCAAGGTCCACCAGGCTTCAGAGGAAACCCAGGTGAGCCTGGTCCTCCAGGCCCTCCAGG
       3390      3400      3410      3420      3430      3440

     3450      3460      3470      3480      3490      3500
ECS_A3 GAGCCACGTGAAAGGCATAAAAGGAGACAAGGGACTCATGGGTGAGCCTGGCCCCAGAGG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Normal GAGCCACGTGAAAGGCATAAAAGGAGACAAGGGACTCATGGGTGAGCCTGGCCCCAGAGG
       3450      3460      3470      3480      3490      3500

     3510      3520      3530      3540      3550      3560
ECS_A3 TCTACCTGGAACTGTGGGAGACAAGGGGCCACCGGGTCAACCGGGAGCACCAGGGAGCCC
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Normal TCTACCTGGAACTGTGGGAGACAAGGGGCCACCGGGTCAACCGGGAGCACCAGGGAGCCC
       3510      3520      3530      3540      3550      3560

     3570      3580      3590      3600      3610      3620
ECS_A3 AGGTCTGCCAGGGCTCAGGGGCGATCCTGGATTCTATGGATTTCCAGGTGTTAAAGGAGA
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Normal AGGTCTGCCAGGGCTCAGGGGCGATCCTGGATTCTATGGATTTCCAGGTGTTAAAGGAGA
       3570      3580      3590      3600      3610      3620

     3630      3640      3650      3660      3670      3680
ECS_A3 GAAGGGTAACCCAGGATTTCCGGGACCAGCTGGACCTCCAGGGCAAATTGGGCCAAAAGG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Normal GAAGGGTAACCCAGGATTTCCGGGACCAGCTGGACCTCCAGGGCAAATTGGGCCAAAAGG
       3630      3640      3650      3660      3670      3680

     3690      3700      3710      3720      3730      3740
ECS_A3 ACCACCTGGTGTCCGTGGAGACCCTGGCACAGTTAAGATCATCTCCCTTCCAGGAAGCCC
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Normal ACCACCTGGTGTCCGTGGAGACCCTGGCACAGTTAAGATCATCTCCCTTCCAGGAAGCCC
       3690      3700      3710      3720      3730      3740

     3750      3760      3770      3780      3790      3800
ECS_A3 AGGGCCACCTGGCCGTGCTGGAGGACCGGGGATGCCAGGAGAACCAGGGCCACCAGGGCC
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Normal AGGGCCACCTGGCCGTGCTGGAGGACCGGGGATGCCAGGAGAACCAGGGCCACCAGGGCC
       3750      3760      3770      3780      3790      3800

     3810      3820      3830      3840      3850      3860
ECS_A3 ACCAGGAATCCTAGGACCCTGTGGACCAAGAGGTAAACCAGGCATGGATGGAATACCAGG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Normal ACCAGGAATCCTAGGACCCTGTGGACCAAGAGGTAAACCAGGCATGGATGGAATACCAGG
       3810      3820      3830      3840      3850      3860

     3870      3880      3890      3900      3910      3920
ECS_A3 AACTCCTGGGCCAATTGGAGAAAAAGGCAACAAAGGTTGTAAAGGCGAGCAAGGACCACC
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Normal AACTCCTGGGCCAATTGGAGAAAAAGGCAACAAAGGTTGTAAAGGCGAGCAAGGACCACC
       3870      3880      3890      3900      3910      3920
```

```
     3930      3940      3950      3960      3970      3980
ECS_A3 TGGATTAGATGGACTGCCAGGCTTGAAGGGGAGACCTGGAGACCCTGGACCACCTGCAAC
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Normal TGGATTAGATGGACTGCCAGGCTTGAAGGGGAGACCTGGAGACCCTGGACCACCTGCAAC
       3930      3940      3950      3960      3970      3980

     3990      4000      4010      4020      4030      4040
ECS_A3 TGGGACCATGAAGAGGGGCTTTATCTTCACCAGGCATAGTCAGACCACAGTGATTCCTTC
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Normal TGGGACCATGAAGAGGGGCTTTATCTTCACCAGGCATAGTCAGACCACAGTGATTCCTTC
       3990      4000      4010      4020      4030      4040

     4050      4060      4070      4080      4090      4100
ECS_A3 CTGTCCAGAAGGGACAGCGCCACTCTATAGTGGCTTTTCTCTTCTTTTCATACAAGGAAA
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Normal CTGTCCAGAAGGGACAGCGCCACTCTATAGTGGCTTTTCTCTTCTTTTCATACAAGGAAA
       4050      4060      4070      4080      4090      4100

     4110      4120      4130      4140      4150      4160
ECS_A3 TGAACAAGCCCATGGACAAGACTTGGGAACTCTCGGCAGCTGCCTGCAGCGATTTACCAC
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Normal TGAACAAGCCCATGGACAAGACTTGGGAACTCTCGGCAGCTGCCTGCAGCGATTTACCAC
       4110      4120      4130      4140      4150      4160

     4170      4180      4190      4200      4210      4220
ECS_A3 AATGCCATTCTTGTTCTGTAACATCAACAATGTCTGTAATTTTGCATCTCGAAATGACTA
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Normal AATGCCATTCTTGTTCTGTAACATCAACAATGTCTGTAATTTTGCATCTCGAAATGACTA
       4170      4180      4190      4200      4210      4220

     4230      4240      4250      4260      4270      4280
ECS_A3 CTCATACTGGCTGTCAACACCAGCTCTGATGCCAATGGACATGGCTCCAATTACTGGCAG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Normal CTCATACTGGCTGTCAACACCAGCTCTGATGCCAATGGACATGGCTCCAATTACTGGCAG
       4230      4240      4250      4260      4270      4280

     4290      4300      4310      4320      4330      4340
ECS_A3 GGCCCTGGAGCCTTATATTAGCAGATGCACTGTCTGTGAAGGTCCTACGATTGCCATAGC
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Normal GGCCCTGGAGCCTTATATTAGCAGATGCACTGTCTGTGAAGGTCCTACGATTGCCATAGC
       4290      4300      4310      4320      4330      4340

     4350      4360      4370      4380      4390      4400
ECS_A3 CATTCACAGCCAAACCACTGACATTCCCTCATGTCCCAATGGCTGGATTTCTCTCTGGAA
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Normal CATTCACAGCCAAACCACTGACATTCCCTCATGTCCCAATGGCTGGATTTCTCTCTGGAA
       4350      4360      4370      4380      4390      4400

     4410      4420      4430      4440      4450      4460
ECS_A3 AGGATTTTCGTTTATCATGTTCACAAGTGCAGGTTCTGAGGGTGCTGGGCAAGCACTGGC
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Normal AGGATTTTCGTTTATCATGTTCACAAGTGCAGGTTCTGAGGGTGCTGGGCAAGCACTGGC
       4410      4420      4430      4440      4450      4460

     4470      4480      4490      4500      4510      4520
ECS_A3 CTCCCCCGGCTCCTGCCTGGAAGAGTTCCGAGCCAGTCCGTTTATASAATGTCACGGAAG
       ::::::::::::::::::::::::::::::::::::::::::::::::.:::::::::::::
Normal CTCCCCCGGCTCCTGCCTGGAAGAGTTCCGAGCCAGTCCGTTTATAGAATGTCACGGAAG
       4470      4480      4490      4500      4510      4520

     4530      4540      4550      4560      4570      4580
ECS_A3 AGGAACGTGCAACTACTATTCAAATTCCTACAGTTTCTGGTTGGCTTCGTTAAACCCCCA
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Normal AGGAACGTGCAACTACTATTCAAATTCCTACAGTTTCTGGTTGGCTTCGTTAAACCCCCA
       4530      4540      4550      4560      4570      4580
```

```
     4590      4600      4610      4620       4630      4640
ECS_A3 AAGAATGTTCAGAAAACCTATTCCATCAACTGTGAAAGCTGAATCCATATGACTAGT
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Normal AAGAATGTTCAGAAAACCTATTCCATCAACTGTGAAAGCTGAATCCATATGACTAGT
       4590      4600      4610      4620      4630      4640
```

APPENDIX 8

Comparison of ECS-HN and normal COL4A4

```
          30        40        50        60        70        80
Normal GGATGTGAGAGTCTTCTGGAAGAGAAGAGAAACCTGACAATCTTTTTTAAGTCCCAAGAC
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 GGATGTGAGAGTCTTCTGGAAGAGAAGAGAAACCTGACAATCTTTTTTAAGTCCCAAGAC
               10        20        30        40        50        60

             90       100       110       120       130       140
Normal TTGACTGTGAGGTAATCCATGCACATGGCATCAGTAAAGTGTTCATTCAGGTGGATCAAG
       :::::::::::::::::::::::::::::::::: :::::::::::::::::::::::::
ESC A4 TTGACTGTGAGGTAATCCATGCACATGGCATCAGTAAGGTGTTCATTCAGGTGGATCAAG
               70        80        90       100       110       120

            150       160       170       180       190       200
Normal CCCTTGGCCAGAGATCCCTGGTCACTTATATTTATCCTCTTTTCTGTACAACATGCATAT
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 CCCTTGGCCAGAGATCCCTGGTCACTTATATTTATCCTCTTTTCTGTACAACATGCATAT
               130       140       150       160       170       180

            210       220       230       240       250       260
Normal GGGAGTGGAAAGAAGTTTGTCGGCCCCTGTGGAGGAAGAGATTGCTCGGTGTGCCATTGC
       :::::::::::: ::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 GGGAGTGGAAAGNAGTTTGTCGGCCCCTGTGGAGGAAGAGATTGCTCGGTGTGCCATTGC
               190       200       210       220       230       240

            270       280       290       300       310       320
Normal TTTCCTGAAAAAGGGTCTCGGGGTCAACCAGGACCGCCGGGGCCACAGGGTCCCATTGGA
        ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 CTTCCTGAAAAAGGGTCTCGGGGTCAACCAGGACCGCCGGGGCCACAGGGTCCCATTGGA
               250       260       270       280       290       300

            330       340       350       360       370       380
Normal CCCCTGGGACTGCCAGGACCCACAGGAATTCCAGGAGAGAAAGGGATGAGAGGTGACAAG
       :::::::::::::::::::::::::::::::::::::::::::::::::::: :::::::::
ESC A4 CCCCTGGGACTGCCAGGACCCACAGGAATTCCAGGAGAGAAAGGGATGAGAAGTGACAAG
               310       320       330       340       350       360

             390       400       410       420       430       440
Normal TGGCCCTCCTGGAGCAGCAGGTGACAAAGGTGACAAGGGTCCAACTGGTGTTCCTGGATT
       :::::::::::::: ::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 TGGCCCTCCTGGAACAGCAGGTGACAAAGGTGACAAGGGTCCAACTGGTGTTCCTGGATT
               370       380       390       400       410

             450       460       470       480       490       500
Normal TCCAGGGTTGGATGGCATACCTGGCCCCCCGGGGCCTCCTGGATCCAGAGGCAAGCCTGG
       :::::::::::::::::::::::::::::::::::::::::::::::::::: ::::::
ESC A4 TCCAGGGTTGGATGGCATACCTGGCCCCCCGGGGCCTCCTGGATCCAGAGGCAACCCTGG
     420       430       440       450       460       470

             510       520       530       540       550       560
Normal CATGCATGGCTACAATGGTTCACGAAGGTGATCCAGGGTTTCCAGGAAAAAGAGGAGTTC
       :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 CATGCATGGCTACAATGGTTCACGAAGGTGATCCAGGGTTTCCAGGAAAAAGAGGAGTTC
     480       490       500       510       520       530

               570       580       590       600       610       620
Normal CTGGCCCAGGAGGCCCCCCAGGCCTTCCTGGGGAAAGGGGAGAAAAAGGAAATTCAGTGT
       :::::::::::::::::::::::::::::: :::::::::::::::::::::::::::::
ESC A4 CTGGCCCAGGAGGCCCCCCAGGCCTTCCTGGGNAAAGGGGAGAAAAAGGAAATTCAGTGT
     540       550       560       570        580        590
```

```
              630       640       650       660       670       680
Normal TCATTTTAGGTGCCATTAAAGGTATTCAGGGAGACAGGGGGGACCCAGACCTGCCTGGCT
       :::::::::::::::::::::::::::::: :::::::::::::::::::: :::::::
ESC A4 TCATTTTAGGTGCCATTAAAGGTATTCAGGGNGACAGGGGGGACCCAGACCTNCCTGGCT
       600       610       620       630       640       650

              690       700       710       720       730       740
Normal TGCCGGGACCGAAGGGGGCAGGAGGACCGGCAGGTCCCAAGGGACAGCCAGGAGAGCCTG
       :::::::::::::::::::: :::::::::::: :::::::::::::::::::::::::
ESC A4 TGCCGGGACCGAAGGGGGCACGAGGACCGGCAGGNCCCAAGGGGACAGCCAGGAGAGCCT
        660       670       680       690       700       710

              750       760       770       780       790       800
Normal GGTTAACAGGTGCTCCGGGCCATCCTGGGAGACCAGGCTTGAAGGGTAATCCTGGTGTGG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 GGTTAACAGGTGCTCCGGGCCATCCTGGGAGACCAGGCTTGAAGGGTAATCCTGGTGTGG
        720       730       740       750       760       770

              810       820       830       840       850       860
Normal GAGTAAAGGGGCAAATGGGAGACCCGGGTGAGGTTGGCCAGCAGGGTTCTCCTGGACCCA
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 GAGTAAAGGGGCAAATGGGAGACCCGGGTGAGGTTGGCCAGCAGGGTTCTCCTGGACCCA
        780       790       800       810       820       830

              870       880       890       900       910       920
Normal CCTTATTGGTACAGCCACCTGATTCCTGTCTGTATAAAGGAGAAAAGGGCATAAAAGGAA
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 CCTTATTGGTACAGCCACCTGATTCCTGTCTGTATAAAGGAGAAAAGGGCATAAAAGGAA
        840       850       860       870       880       890

              930       940       950       960       970       980
Normal TGCCTGGTATGATTGGACCTCCAGGACTACCGGGACCCAAGGGAGAACCTGGAATTGGAC
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 TGCCTGGTATGATTGGACCTCCAGGACTACCGGGACCCAAGGGAGAACCTGGAATTGGAC
        900       910       920       930       940       950

              990      1000      1010      1020      1030      1040
Normal CAAAAGGAGAGAAGGGTATTCCTGGGTTCTCAGGACCTCGGGGTGATCCGGGTTCCTATG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 CAAAAGGAGAGAAGGGTATTCCTGGGTTCTCAGGACCTCGGGGTGATCCGGGTTCCTATG
        960       970       980       990      1000      1010

             1050      1060      1070      1080      1090      1100
Normal GCTCTCCAGGTTTTCCAGGATTAAAGGGGAAACCAGGACTGTTTGGAGATCCTGGATCAT
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 GCTCTCCAGGTTTTCCAGGATTAAAGGGGAAACCAGGACTGTTTGGAGATCCTGGATCAT
       1020      1030      1040      1050      1060      1070

             1110      1120      1130      1140      1150      1160
Normal TTGGATTTGTTGGCCCAAAGGGGGATCCTGGAGACCGTGGGTACCCAGGACCACCGGGTG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 TTGGATTTGTTGGCCCAAAGGGGGATCCTGGAGACCGTGGGTACCCAGGACCACCGGGTG
       1080      1090      1100      1110      1120      1130

             1170      1180      1190      1200      1210      1220
Normal TTCTGGTAACTCCATCTGTTCCACTCAAAGGCCCTCCAGGGGATCCGGGGCGCCCTGGCC
       :::::::::::::: ::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 TTCTGGTAACTCCACCTGTTCCACTCAAAGGCCCTCCAGGGGATCCGGGGCGCCCTGGCC
       1140      1150      1160      1170      1180      1190

             1230      1240      1250      1260      1270      1280
Normal ACTATGGAGAAATGGGGTCCGTTGGACCACCTGGTCCCCCTGGTCCCCCAGGTCCACCAG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 ACTATGGAGAAATGGGGTCCGTTGGACCACCTGGTCCCCCTGGTCCCCCAGGTCCACCAG
       1200      1210      1220      1230      1240      1250
```

```
             1290      1300      1310      1320      1330      1340
Normal GGGAAGACTGTGCAGGCATGATGGGACCCCCTGGACCAAGAGGATTTCCTGGTCATCCAG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 GGGAAGACTGTGCAGGCATGATGGGACCCCCTGGACCAAGAGGATTTCCTGGTCATCCAG
       1260      1270      1280      1290      1300      1310

             1350      1360      1370      1380      1390      1400
Normal GATTTCCAGGGGCAGCTGGTATCCCAGGGAGAGCGGATTTAAGTCCAGGAAAACCAGGGA
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 GATTTCCAGGGGCAGCTGGTATCCCAGGGAGAGCGGATTTAAGTCCAGGAAAACCAGGGA
       1320      1330      1340      1350      1360      1370

             1410      1420      1430      1440      1450      1460
Normal AACCGGGACCTCCTGGGTTGCCCGGAGCACCAGGGCTGCAAGGACCCCCAGGATCAGATG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 AACCGGGACCTCCTGGGTTGCCCGGAGCACCAGGGCTGCAAGGACCCCCAGGATCAGATG
       1380      1390      1400      1410      1420      1430

             1470      1480      1490      1500      1510      1520
Normal TTATCTATTGTAGTGTTGGGCATCCTGGACCACAAGGGATAAAAGGCAAAATGGGTCCTC
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 TTATCTATTGTAGTGTTGGGCATCCTGGACCACAAGGGATAAAAGGCAAAATGGGTCCTC
       1440      1450      1460      1470      1480      1490

             1530      1540      1550      1560      1570      1580
Normal CAGGAAGAAGAGGCTCGAAAGGAGAAAAAGGAAATGCGGGGCTCTGTGCCTGTGAGCCTG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 CAGGAAGAAGAGGCTCGAAAGGAGAAAAAGGAAATGCGGGGCTCTGTGCCTGTGAGCCTG
       1500      1510      1520      1530      1540      1550

             1590      1600      1610      1620      1630      1640
Normal GTCCCATGGGCCCACCAGGCCCTCCAGGACTTCCTGGAAGGCAGGGTAGTAAGGGAGACT
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 GTCCCATGGGCCCACCAGGCCCTCCAGGACTTCCTGGAAGGCAGGGTAGTAAGGGAGACT
       1560      1570      1580      1590      1600      1610

             1650      1660      1670      1680      1690      1700
Normal TGGGGCTCCCTGGGTGGATTGGAGAGAAAGGTCACCCAGGCCCTCCTGGTGCTGAAGGAT
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 TGGGGCTCCCTGGGTGGATTGGAGAGAAAGGTCACCCAGGCCCTCCTGGTGCTGAAGGAT
       1620      1630      1640      1650      1660      1670

             1710      1720      1730      1740      1750      1760
Normal CTCCAGGATCACCAGGAAAACATGGCGCCTCAGGACCACCTGGCAGCAAAGGAGAAAAGG
       :::::::::::::::::::::::: ::::::::::::::::::::::::::::::::::
ESC A4 CTCCAGGATCACCAGGAAAACATGGTGCCTCAGGACCACCTGGCAGCAAAGGAGAAAAGG
       1680      1690      1700      1710      1720      1730

             1770      1780      1790      1800      1810      1820
Normal GCGACAGGGTTGTATCAAGAGTGAAAGGGAAAAAAGGAGAAAGAGGTCCTGATGGGCTCC
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 GCGACAGGGTTGTATCAAGAGTGAAAGGGAAAAAAGGAGAAAGAGGTCCTGATGGGCTCC
       1740      1750      1760      1770      1780      1790

             1830      1840      1850      1860      1870      1880
Normal CAGGATTTCCAGGACAACAGGGACAAAATGGTCGGGATGGACTTCCTGGAAAAAAGGGGG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 CAGGATTTCCAGGACAACAGGGACAAAATGGTCGGGATGGACTTCCTGGAAAAAAGGGGG
       1800      1810      1820      1830      1840      1850

             1890      1900      1910      1920      1930      1940
Normal ATCCAGGCCCCCCGGGGGGTCATGAGGATGCAGCCCCAGGTGCTAAAGGGTCTCCTGGAC
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 ATCCAGGCCCCCCGGGGGGTCATGAGGATGCAGCCCCAGGTGCTAAAGGGTCTCCTGGAC
       1860      1870      1880      1890      1900      1910
```

```
             1950      1960      1970      1980      1990      2000
Normal CACCGGGGCCCCCAGGCAGAACAGGACCCAGGGGACAACCAGGTCTGGGATTTCCTGGTC
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 CACCGGGGCCCCCAGGCAGAACAGGACCCAGGGGACAACCAGGTCTGGGATTTCCTGGTC
       1920      1930      1940      1950      1960      1970

           2010      2020      2030      2040      2050      2060
Normal CACCAGGAGAGAGAGGACAACCAGGAGCTCCAGGCCGCCCTGGCGAGAGGGGCTTCGGGG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 CACCAGGAGAGAGAGGACAACCAGGAGCTCCAGGCCGCCCTGGCGAGAGGGGCTTCGGGG
       1980      1990      2000      2010      2020      2030

             2070      2080      2090      2100      2110      2120
Normal GCTTGAAGGGCCAGAAAGGTGATACAATTTCTTGTAATGTCACCTACCCTGGGAGGCCAG
       :::::::::::::::::::::::::::::::::: :::::::::::::::::::: :
ESC A4 GCTTGAAGGGCCAGAAAGGTGATACAATTTCTTGTAACGTCACCTACCCTGGGAGGCCGG
       2040      2050      2060      2070      2080      2090

             2130      2140      2150      2160      2170      2180
Normal GACCCCCAGGTTTTGATGGACCTCCAGGACCAAAGGGATTTCCAGGTCCTCAAGGCGCTC
       : ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 GGCCCCCAGGTTTTGATGGACCTCCAGGACCAAAGGGATTTCCAGGTCCTCAAGGCGCTC
       2100      2110      2120      2130      2140      2150

             2190      2200      2210      2220      2230      2240
Normal CAGGGTTGAGGTGCTTGGATGGGCCAAAAGGTCAGCGTGGCAAACCAGGAATGTCAGAGA
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 CAGGGTTGAGGTGCTTGGATGGGCCAAAAGGTCAGCGTGGCAAACCAGGAATGTCAGAGA
       2160      2170      2180      2190      2200      2210

             2250      2260      2270      2280      2290      2300
Normal TACCCGGTCCACCTGGTTTTCGTGGTGACATGGGAGATCCGGGTTTTGAAGGTGAAAAGG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 TACCCGGTCCACCTGGTTTTCGTGGTGACATGGGAGATCCGGGTTTTGAAGGTGAAAAGG
       2220      2230      2240      2250      2260      2270

             2310      2320      2330      2340      2350      2360
Normal GGTCCTCCCTTCTTGGGCCGCCAGGCTTTCCTGGTTCTCGTGGAGCAAATGGTCAGAAAG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 GGTCCTCCCTTCTTGGGCCGCCAGGCTTTCCTGGTTCTCGTGGAGCAAATGGTCAGAAAG
       2280      2290      2300      2310      2320      2330

             2370      2380      2390      2400      2410      2420
Normal GAATGATGGGAGACATTGCCTATGGCCTCCCAGGCCCCCCAGGAAAGAGAGGTCCTTCCG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 GAATGATGGGAGACATTGCCTATGGCCTCCCAGGCCCCCCAGGAAAGAGAGGTCCTTCCG
       2340      2350      2360      2370      2380      2390

             2430      2440      2450      2460      2470      2480
Normal GAGTGCCAGGGTCCAAAGGACACAGAGGTGATTCAGGACGTCCAGGGTTTGCAGGGCCAG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 GAGTGCCAGGGTCCAAAGGACACAGAGGTGATTCAGGACGTCCAGGGTTTGCAGGGCCAG
       2400      2410      2420      2430      2440      2450

             2490      2500      2510      2520      2530      2540
Normal CCGGCAAGCCTGGATCCCCAGGTCTCAAAGGTCCCAGAGGCAGAGAGGGAAGTGCTGGGT
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 CCGGCAAGCCTGGATCCCCAGGTCTCAAAGGTCCCAGAGGCAGAGAGGGAAGTGCTGGGT
       2460      2470      2480      2490      2500      2510

             2550      2560      2570      2580      2590      2600
Normal TTCAAGGTATCCCGGGTCCACCTGGCCATTCCTGCAAAAGAGGCGCTCCAGGGATACCAG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 TTCAAGGTATCCCGGGTCCACCTGGCCATTCCTGCAAAAGAGGCGCTCCAGGGATACCAG
       2520      2530      2540      2550      2560      2570
```

```
             2610      2620      2630      2640      2650      2660
Normal GGCAACCGGGGCTCCCTGGGGCCCCGGGGAGTCCAGGTGCCCCAGGTTGGAAAGGACAGC
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 GGCAACCGGGGCTCCCTGGGGCCCCGGGGAGTCCAGGTGCCCCAGGTTGGAAAGGACAGC
       2580      2590      2600      2610      2620      2630

             2670      2680      2690      2700      2710      2720
Normal GAGGGGATGTGGGGCCTCCGGGTCCAGCTGGAATGAAAGGCCTCCCTGGAGTCCCTGGAC
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 GAGGGGATGTGGGGCCTCCGGGTCCAGCTGGAATGAAAGGCCTCCCTGGAGTCCCTGGAC
       2640      2650      2660      2670      2680      2690

             2730      2740      2750      2760      2770      2780
Normal GGCCAGGGGCAGAGGGACCCCTAGGATTCCCAGGAGTCCCAGGCCCCTTAGGGGATGATG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 GGCCAGGGGCAGAGGGACCCCTAGGATTCCCAGGAGTCCCAGGCCCCTTAGGGGATGATG
       2700      2710      2720      2730      2740      2750

             2790      2800      2810      2820      2830      2840
Normal GGCTGCCTGGTCTTCCAGGCCCGAAGGGACCCCAGGGGCTGCCTGGCTTCCCCGGTTTTC
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 GGCTGCCTGGTCTTCCAGGCCCGAAGGGACCCCAGGGGCTGCCTGGCTTCCCCGGTTTTC
       2760      2770      2780      2790      2800      2810

             2850      2860      2870      2880      2890      2900
Normal CAGGGGAGAGAGGAAAGCCTGGCCCAGAGGGACGCCCTGGCAGCAAGGGGGAACCAGGAG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 CAGGGGAGAGAGGAAAGCCTGGCCCAGAGGGACGCCCTGGCAGCAAGGGGGAACCAGGAG
       2820      2830      2840      2850      2860      2870

             2910      2920      2930      2940      2950      2960
Normal AGGATGGTCGGCCTGGCTTCTTCGGAGACCAAGGGGTGAAAGGTGCCAAAGGAGCGAGAG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 AGGATGGTCGGCCTGGCTTCTTCGGAGACCAAGGGGTGAAAGGTGCCAAAGGAGCGAGAG
       2880      2890      2900      2910      2920      2930

             2970      2980      2990      3000      3010      3020
Normal GACCCCCAGGAGATGAAGGAGAGATGGCGGTCTTTTCCCAAAAAGGGAAAACCGGGGAAC
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 GACCCCCAGGAGATGAAGGAGAGATGGCGGTCTTTTCCCAAAAAGGGAAAACCGGGGAAC
       2940      2950      2960      2970      2980      2990

             3030      3040      3050      3060      3070      3080
Normal CTGGACCTCCAGGAGATGATGGATTCCCAGGAGAAGAAGGTGATAAAGGCAATCCAGGGA
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 CTGGACCTCCAGGAGATGATGGATTCCCAGGAGAAGAAGGTGATAAAGGCAATCCAGGGA
       3000      3010      3020      3030      3040      3050

             3090      3100      3110      3120      3130      3140
Normal TGCCGGGGAGGAGAGGAGAGCCTGGAAGATTCGGAGCACCTGGATTTCACAGAGGGGAGC
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 TGCCGGGGAGGAGAGGAGAGCCTGGAAGATTCGGAGCACCTGGATTTCACAGAGGGGAGC
       3060      3070      3080      3090      3100      3110

             3150      3160      3170      3180      3190      3200
Normal CCGGGAGAACCGGGCAGCCAGGGCTTCCTGGCCCCCCAGGCCCCCCAGGCTCACCTGGGC
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 CCGGGAGAACCGGGCAGCCAGGGCTTCCTGGCCCCCCAGGCCCCCCAGGCTCACCTGGGC
       3120      3130      3140      3150      3160      3170

             3210      3220      3230      3240      3250      3260
Normal TGAGAGGGATTATTGGTTTTCCGGGATTTCCAGGTGACCAGGGTGAGCCAGGTTCTCCAG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 TGAGAGGGATTATTGGTTTTCCGGGATTTCCAGGTGACCAGGGTGAGCCAGGTTCTCCAG
       3180      3190      3200      3210      3220      3230
```

```
             3270      3280      3290      3300      3310      3320
Normal GGTCTCCTGGACTTTCAGGAATTGATGGAATGAGAGGACCTAAAGGAAGCAAAGGTGACC
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 GGTCTCCTGGACTTTCAGGAATTGATGGAATGAGAGGACCTAAAGGAAGCAAAGGTGACC
       3240      3250      3260      3270      3280      3290

             3330      3340      3350      3360      3370      3380
Normal CTGCAAGTCAGTTCGGCCCACCTGGTCCAAAGGGTGAACCAGGTAGCCCTGGATGTCCAG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 CTGCAAGTCAGTTCGGCCCACCTGGTCCAAAGGGTGAACCAGGTAGCCCTGGATGTCCAG
       3300      3310      3320      3330      3340      3350

              3390      3400      3410      3420      3430      3440
Normal GACAAGTTGGAGCACCTGGAGAGCAGGGCTTGCCTGGTGTTCAAGGGCTCAGAGGACCAC
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 GACAAGTTGGAGCACCTGGAGAGCAGGGCTTGCCTGGTGTTCAAGGGCTCAGAGGACCAC
       3360      3370      3380      3390      3400      3410

             3450      3460      3470      3480      3490      3500
Normal CTGGAAGGCCAGGATTACCTGGCTCCTCCGGACCACCAGGGTGTCCAGGTAATCAAGGGG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 CTGGAAGGCCAGGATTACCTGGCTCCTCCGGACCACCAGGGTGTCCAGGTAATCAAGGGG
       3420      3430      3440      3450      3460      3470

             3510      3520      3530      3540      3550      3560
Normal TGCCTGGGCTGACAGGGCCTCCAGGAGAAACAGGGGATCCTGGGCCAAGAGGCATGATGG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 TGCCTGGGCTGACAGGGCCTCCAGGAGAAACAGGGGATCCTGGGCCAAGAGGCATGATGG
       3480      3490      3500      3510      3520      3530

             3570      3580      3590      3600      3610      3620
Normal GAGATCCAGGGCCACCAGGTCTTCCAGGAATAAAAGGTCCCTCCGGGTCGCCGGGTCTGA
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 GAGATCCAGGGCCACCAGGTCTTCCAGGAATAAAAGGTCCCTCCGGGTCGCCGGGTCTGA
       3540      3550      3560      3570      3580      3590

             3630      3640      3650      3660      3670      3680
Normal ATGGCTTGCATGGTTTAAAGGGTCAGAAAGGAGCCAAAGGTGCTTCAGGTTTGCACGAAG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 ATGGCTTGCATGGTTTAAAGGGTCAGAAAGGAGCCAAAGGTGCTTCAGGTTTGCACGAAG
       3600      3610      3620      3630      3640      3650

             3690      3700      3710      3720      3730      3740
Normal TGGGCCCACCCGGTCCAGTGGGCGTACCTGGGCTGAAAGGGGAGCCGGGAGACCCTGGGA
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 TGGGCCCACCCGGTCCAGTGGGCGTACCTGGGCTGAAAGGGGAGCCGGGAGACCCTGGGA
       3660      3670      3680      3690      3700      3710

             3750      3760      3770      3780      3790      3800
Normal GCCCAGGAATTTCTCCCCCAGGCCTTTCTGGAGAAAGAGGCCCCCCCGGCCCCCCAGGGA
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 GCCCAGGAATTTCTCCCCCAGGCCTTTCTGGAGAAAGAGGCCCCCCCGGCCCCCCAGGGA
       3720      3730      3740      3750      3760      3770

             3810      3820      3830      3840      3850      3860
Normal GACCTGGAGCCCCTGGTCCTGCAGGTGCCACAGGAAGAGCTGCTGAAGGTCACGTTCCTG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 GACCTGGAGCCCCTGGTCCTGCAGGTGCCACAGGAAGAGCTGCTGAAGGTCACGTTCCTG
       3780      3790      3800      3810      3820      3830

             3870      3880      3890      3900      3910      3920
Normal ACCCAGGTCCACCCGGAGATGTGGGTCCTCCTGGCCCCGATGGTCCGAGAGGAGCGCCCG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 ACCCAGGTCCACCCGGAGATGTGGGTCCTCCTGGCCCCGATGGTCCGAGAGGAGCGCCCG
       3840      3850      3860      3870      3880      3890
```

```
             3930      3940      3950      3960      3970      3980
Normal GGCCCCCAGGCCCCCCTGGGAGCGTTGACCTTCTGAGAGGGGAACCAGGAGACTGTGGGC
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 GGCCCCCAGGCCCCCCTGGGAGCGTTGACCTTCTGAGAGGGGAACCAGGAGACTGTGGGC
       3900      3910      3920      3930      3940      3950

             3990      4000      4010      4020      4030      4040
Normal CGCCGGGGCCTCCAGGTCCCCCGGGCCCACCCGGCCCTCCAGGACACCAAGGCTTCCCAG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 CGCCGGGGCCTCCAGGTCCCCCGGGCCCACCCGGCCCTCCAGGACACCAAGGCTTCCCAG
       3960      3970      3980      3990      4000      4010

             4050      4060      4070      4080      4090      4100
Normal GATGCGATGGAAAAGACGGCCAGAAAGGACCAATAGGATTCCCGGGGCTGCAGGGGCCTC
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 GATGCGATGGAAAAGACGGCCAGAAAGGACCAATAGGATTCCCGGGGCTGCAGGGGCCTC
       4020      4030      4040      4050      4060      4070

             4110      4120      4130      4140      4150      4160
Normal AGGGACTTCCTGGGCTCCCTGGGGAGAAGGGTCTGCTGGGCATTCCAGGCCGGCGAGGGC
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 AGGGACTTCCTGGGCTCCCTGGGGAGAAGGGTCTGCTGGGCATTCCAGGCCGGCGAGGGC
       4080      4090      4100      4110      4120      4130

             4170      4180      4190      4200      4210      4220
Normal ACCCCGGTCCTCCAGGTTCCAGAGGTGAACCAGGGCCTCCTGCAGATGTGGATTCGTGCC
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 ACCCCGGTCCTCCAGGTTCCAGAGGTGAACCAGGGCCTCCTGCAGATGTGGATTCGTGCC
       4140      4150      4160      4170      4180      4190

             4230      4240      4250      4260      4270      4280
Normal CCCGCATCCCCGGGCTTCCTGGGGTACCAGGCCCAAGAGGACCAGAAGGAGCCATGGGGG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 CCCGCATCCCCGGGCTTCCTGGGGTACCAGGCCCAAGAGGACCAGAAGGAGCCATGGGGG
       4200      4210      4220      4230      4240      4250

             4290      4300      4310      4320      4330      4340
Normal TCCCTGGAAGGAGAGGGCCCCCAGGACCAGGGTGCAAAGGAGAGCCGGGACTGGAGGGCA
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 TCCCTGGAAGGAGAGGGCCCCCAGGACCAGGGTGCAAAGGAGAGCCGGGACTGGAGGGCA
       4260      4270      4280      4290      4300      4310

             4350      4360      4370      4380      4390      4400
Normal GGAGGGGCGAGGCCGGCCTCCCAGGACCTCCGGGGCCTCCCGGAAGCACAGGGGACGCAG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 GGAGGGGCGAGGCCGGCCTCCCAGGACCTCCGGGGCCTCCCGGAAGCACAGGGGACGCAG
       4320      4330      4340      4350      4360      4370

             4410      4420      4430      4440      4450      4460
Normal GAGAAGCCGGTTGCCCCGGAGCACCAGGGCCCCCCGGGCCCATTGGGGACCCCGGGCCCA
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 GAGAAGCCGGTTGCCCCGGAGCACCAGGGCCCCCCGGGCCCATTGGGGACCCCGGGCCCA
       4380      4390      4400      4410      4420      4430

             4470      4480      4490      4500      4510      4520
Normal GAGGGTTTGGGCCTGGATACCTCAGTGGCTTCCTCCTGGTTCTCCACAGTCAGACGGATG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 GAGGGTTTGGGCCTGGATACCTCAGTGGCTTCCTCCTGGTTCTCCACAGTCAGACGGATG
       4440      4450      4460      4470      4480      4490

             4530      4540      4550      4560      4570      4580
Normal GGGAGCCCACCTGCCCCGCGGGCATGCCCAGGCCCTGGACGGGCTACAGTCTGTTATACC
       :::::::::::::::::::::::::::::::: ::::::::::::::::::::::::::
ESC A4 GGGAGCCCACCTGCCCCGCGGGCATGCCCAGGCTCTGGACGGGCTACAGTCTGTTATACC
       4500      4510       4520      4530      4540      4550
```

```
             4590      4600      4610      4620      4630      4640
Normal TGGAAGGACAGGAGAAGGCCCACAATCAGGACCTTGGTCTGGCAGGGTCTTGCCTTCCCA
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 TGGAAGGACAGGAGAAGGCCCACAATCAGGACCTTGGTCTGGCAGGGTCTTGCCTTCCCA
        4560      4570      4580      4590      4600      4610

             4650      4660      4670      4680      4690      4700
Normal TGTTCAGCACCCTGCCCTTTGCCTACTGCAACATCCACCAAGTGTGCCACTACGCCCGGA
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 TGTTCAGCACCCTGCCCTTTGCCTACTGCAACATCCACCAAGTGTGCCACTACGCCCGGA
        4620      4630      4640      4650      4660      4670

             4710      4720      4730      4740      4750      4760
Normal GAAACGACCGGTCCTACTGGCTGGCCAGTGCTGCGCCCCTGCCCATGATGCCGCTCTCAG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 GAAACGACCGGTCCTACTGGCTGGCCAGTSCTGCGCCCCTGCCCATGATGCCGCTCTCAG
        4680      4690      4700      4710      4720      4730


         4770      4780      4790      4800      4810      4820
Normal AGGAGGAGATCCGCCCCTACATCAGCCGCTGTGCCGTGTGCGAGGCCCCAGCCCCCGTGG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 AGGAGGAGATCCGCCCCTACATCAGCCGCTGTGCCGTGTGCGAGGCCCCAGCCCCCGTGG
        4740      4750      4760      4770      4780      4790

             4830      4840      4850      4860      4870      4880
Normal TGGCGCTGCACAGCCAGGACCGGTCCATCCCGCCGTGTCCACGGAGCTGGAGGAGCCTCT
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 TGGCGCTGCACAGCCAGGACCGGTCCATCCCGCCGTGTCCACGGAGCTGGAGGAGCCTCT
          4800      4810      4820      4830      4840      4850

             4890      4900      4910      4920      4930      4940
Normal GGATCGGGTACTCGTTCCTGATGCACACAGGGGCTGGAGACCAAGGAGGAGGGCAGGCCC
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 GGATCGGGTACTCGTTCCTGATGCACACAGGGGCTGGAGACCAAGGAGGAGGGCAGGCCC
          4860      4870      4880      4890      4900      4910

             4950      4960      4970      4980      4990      5000
Normal TCATGTCCCCCGGCAGCTGTCTGGAAGATTTCCGAGCCGCACCGTTCCTCGAATGCCAAG
       :::::::::::::::::::::::::::::::::::::::::::::::::: ::::::
ESC A4 TCATGTCCCCCGGCAGCTGTCTGGAAGATTTCCGAGCCGCACCGTTCCTCGAACGCCAAG
          4920      4930      4940      4950      4960      4970

             5010      5020      5030           5040               5050
Normal GCCGACAGGGAACTTGCCACTTTTTTGCAAATGAGTATAGCTTCTGGCTTCTGGCTGACGAC
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
ESC A4 GCCGACAGGGAACTTGCCACTTTTTTGCAAATGAGTATAGCTTCTGGCTTCTGGCTGACGAC
          4980      4990      5000      5010              5020
```

APPENDIX 9

Effects of Aberrant COL4A3 Nucleotides

Position	Change	Effect
ECS 225	R	All bases conserve proline (CCN)
ECS 440	Y	All bases conserve leucine (CTN)
ECS 2340	R	All bases conserve valine (GTN)
ECS 2507	A-G	Conserves codon for glutamic acid
ECS 2575	G-A	Conserves codon for glutamic acid
ECS 3250	A-G	Conserves codon for glutamine
ECS 3338	Y	All bases conserve threonine (ACN)
ECS 3220	G-T	serine-isoleucine
ECS 3250	G-A	arginine-lysine
ECS 3254	C-A	Conserves codon for glycine

APPENDIX 10

Effects of Aberrant COL4A4 Nucleotides

Position	Change	Effect
ECS 99	G-A	Conserves AAA codon for lysine
ECS 195	N	Unknown
ECS 351	G-A	Conserves GAA codon for glutamine
ECS 371	G-A	Conserves GAA codon for glutamine
ECS 475	G-C	Asparagine-Lysine
ECS 573	N	Unknown
ECS 630	N	Unknown
ECS 650	N	All bases conserve leucine (CTN)
ECS 691	N	All bases conserve glycine (GGN)
ECS 1179	C-T	Conserves codon for leucine
ECS 1707	C-T	Conserves codon for glycine
ECS 2078	C-T	Conserves codon for asparagine
ECS 2099	C-T	Conserves codon for proline
ECS 2100	A-G	Conserves codon for glycine
ECS 4530	C-T	Conserves codon for glycine

VITA

Name:
Zenaido Tres Camacho

Permanent address:
9393 Carriage Way, College Station, TX 77845

Educational background:

The University of Texas at Austin B.S. 1987
Majour subject: Zoology

Texas A&M University Ph.D. 2003
Majour subject: Veterinary Microbiology